

Федеральное государственное автономное образовательное
учреждение высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

Институт космических и информационных технологий

Кафедра «Информационные системы»

УТВЕРЖДАЮ
Зав. кафедрой ИС
_____ С.А. Виденин
« ____ » _____ 2016 г.

ДИПЛОМНЫЙ ПРОЕКТ

230102.65 Автоматизированные системы обработки информации и
управления

Автоматизация параметрического анализа клеточных структур деревьев
хвойных пород

Пояснительная записка

Руководитель

подпись, дата

П. П. Дьячук
инициалы, фамилия

Выпускник

подпись, дата

А. Ю. Клёнкин
инициалы, фамилия

Нормоконтролер

подпись, дата

Л. С. Троценко
инициалы, фамилия

Красноярск 2016

РЕФЕРАТ

Пояснительная записка к дипломному проекту по теме «Автоматизация параметрического анализа клеточных структур деревьев хвойных пород» представлена на 75 страницах, содержит 32 иллюстрации, 23 формулы, 9 источников используемой литературы.

КЛЕТОЧНЫЕ СТРУКТУРЫ ДЕРЕВЬЕВ ХВОЙНЫХ ПОРОД, ОБРАБОТКА ИЗОБРАЖЕНИЙ, РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ, СЕГМЕНТАЦИЯ ИЗОБРАЖЕНИЙ, АВТОМАТИЗИРОВАННЫЙ КЛЕТОЧНЫЙ АНАЛИЗ, ПАРАМЕТРИЧЕСКИЙ АНАЛИЗ

Объектом исследования является параметрический анализ клеточных структур деревьев хвойных пород.

Предметом исследования являются методы обработки и распознавания фотографических изображений с последующим анализом характеристик выделенных объектов.

Целью исследования является разработка автоматизированной информационной системы параметрического анализа клеточных структур деревьев хвойных пород.

Задачи исследования: анализ предметной области, изучение существующих решений в области анализа клеточных структур, изучение методов обработки и распознавания изображений, разработка информационной системы автоматизированного параметрического анализа деревьев хвойных пород.

Актуальность работы заключается в автоматизации исследовательских функций анализа массивов фотоизображений световой микроскопии и выделению параметров клеточных структур деревьев хвойных пород. В настоящее время указанные функции выполняются в ручном или полуавтоматическом режимах, что существенно влияет на продолжительность и трудоёмкость накопления параметрических данных о клеточных структурах хвойных и, соответственно, сужает возможности их статистической обработки. Разрабатываемая АИС позволяет существенно сократить время на обработку данных и значительно ускорить накопление массивов статистической информации для последующей работы с ними специалистов в области дендрозологии и дендрохронологии.

						ДП-230102.65-001019120ПЗ		
Изм.	Коп.уч	Лист	№док	Подп.	Дата			
Разраб.		Клёнкин А. Ю.			14.06.16	Автоматизация параметрического анализа клеточных структур деревьев хвойных пород	Стадия	Лист
Пров.		Дьячук П. П.			15.06.16			2
Н. контр.		Троценко Л. С.			17.06.16		Кафедра «Информационные системы»	
Утв.		Виденин С. А.			23.06.16			
							Листов	75

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	5
1 Обзор предметной области	7
1.1 Общие сведения о клеточных структурах хвойных	7
1.1.1 Гистометрические измерения клеточных структур хвойных	8
1.1.2 Другие методы измерения клеточных структур хвойных	10
1.2 Применяемая методика измерения структур клеток хвойных	11
2 Обзор существующих решений в области анализа клеток.....	14
2.1 Решения в области анализа клеточных структур крови.....	16
2.1.1 Проточная цитометрия	17
2.1.2 Компьютерный гематологический анализ	18
2.2 Решения в области распознавания и анализа клеточных структур деревьев хвойных пород.....	21
2.2.1 Программа «SuperMoment».....	24
2.2.2 Программа «Lineyka»	24
3 Анализ методов обработки и распознавания изображений.....	28
3.1 Методы повышения контраста изображения	28
3.2 Методы фильтрации изображения.....	31
3.2.1 Методы линейной фильтрации.....	32
3.2.2 Методы нелинейной фильтрации.....	33
3.2.3 Комбинированные и гибридные методы фильтрации	34
3.2.4 Адаптивные методы фильтрации	35
3.3 Методы бинаризации изображения	36
3.3.1 Методы локальной бинаризации	37
3.3.2 Методы глобальной бинаризации	39
4 Реализация системы параметрического анализа клеточных структур деревьев хвойных пород.....	41
4.1 Описание входных данных системы	41
4.2 Состав модулей системы	43
4.3 Предварительная обработка изображения	44
4.3.1 Загрузка изображений	44
4.3.2 Повышение контрастности изображений.....	46
4.3.3 Медианная фильтрация	48
4.3.4 Бинаризация.....	50
4.3.5 Удаление нецелых объектов	54
4.3.6 Выделение объектов	56
4.3.7 Морфологическое закрытие объектов	60
4.4 Обработка массивов распознанных объектов	61
4.4.1 Вычисление центров масс объектов	61

4.4.2	Определение периметра объектов.....	62
4.4.3	Исключение ложных объектов	63
4.4.4	Заполнение полостей объектов.....	68
4.4.5	Исключение пересекающихся точек объектов	69
4.4.6	Вычисление угла ориентации главной оси изображения	71
4.5	Выходные результаты работы системы	72
ЗАКЛЮЧЕНИЕ		74
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ		75
ПРИЛОЖЕНИЕ А Техническое задание		76

Изм.	Колич.	Лист.	№ док	Подпись	Дата

ДП-230102.65-001019120ПЗ

Лист

4

ВВЕДЕНИЕ

Информационные технологии на сегодняшний день - одна из наиболее динамично развивающихся областей человеческой деятельности и их значение в обозримой перспективе будет лишь расти. Это обусловлено тем, что характерной чертой современного мира является лавинообразное увеличение числа информационных потоков, сколь либо адекватно обрабатывать которые без создания всевозможных информационных систем становится невозможно. Именно информационные системы позволяют вывести задачи хранения, поиска, обработки и даже получения информации на недоступный ещё недавно уровень как по масштабу, так и по глубине.

Одними из наиболее важных сфер приложения информационных систем являются автоматизированная обработка изображений и распознавание визуальных образов. Подобная проблематика актуальна для множества прикладных задач: системы безопасности, биологические и особенно медицинские исследования, промышленное конвейерное оборудование, архивное дело и т.д. Задачи распознавания как правило приходится решать в условиях ограниченного временного ресурса, а в некоторых случаях и вовсе в режиме реального времени. Толчком к бурному развитию данного направления стали значительно возросшие за последние 2-3 десятилетия характеристики вычислительной техники и её компонентной базы, позволившие создать и реализовать алгоритмы, взявшие на себя целый класс задач, прежде выполнявшихся преимущественно вручную, а то и вовсе не ставившихся из-за неприемлемых объёмов и сроков выполнения. На сегодняшний день задача получения качественных изображений биологических объектов не представляет научного интереса по причине широкого распространения цифровых фото-, видеосистем, и особое внимание уделяется методам их математической обработки и анализа.

В настоящий момент одной из актуальных проблем в дендрохронологии является параметрический анализ клеточных структур деревьев хвойных пород. Кроме того, понимание закономерностей скорости и ритма роста деревьев на определённой территории, а также факторов, влияющих на них, может также иметь важное значение и в практике лесного хозяйства. Ручная обработка гигантских массивов цитологических фотоизображений срезов древесины малоэффективна и потому получить статистически надёжные результаты от такой деятельности крайне сложно. Перед исследователями встала задача автоматизации параметрического анализа фотоизображений годичных колец, которая активно решается в том числе и учёными Сибирского федерального университета.

В рамках данного дипломного проекта будет разработан прототип автоматизированной системы параметрического анализа клеточных структур деревьев хвойных пород. Чтобы спроектировать данную систему,

необходимо представлять, какие функции она должна выполнять и какие методы необходимо для этого применить. Исходя из этого, в рамках настоящего работы будут рассмотрены предметная область, включая обзор основных принципов проведения параметрического анализа клеточных структур деревьев хвойных пород, существующие решения в области параметрического анализа цитологических объектов, а также методы обработки и распознавания изображений. Результатом указанных теоретических изысканий должно стать описание разработанного прототипа программного продукта для автоматизации задач предметной области исследования.

1 Обзор предметной области

1.1 Общие сведения о клеточных структурах хвойных

Хвойные деревья относятся к классу хвойных отдела голосеменных и включают в себя 7 семейств, около 55 родов и 600 видов. Представители хвойных пород - это, чаще всего, деревья 10 до 100 м в высоту, значительно реже - кустарники. Большинство хвойных деревьев являются вечнозелёными, за исключением некоторых, сбрасывающих хвою на зиму, например, лиственницы, являющейся одной из основных лесообразующих пород РФ.

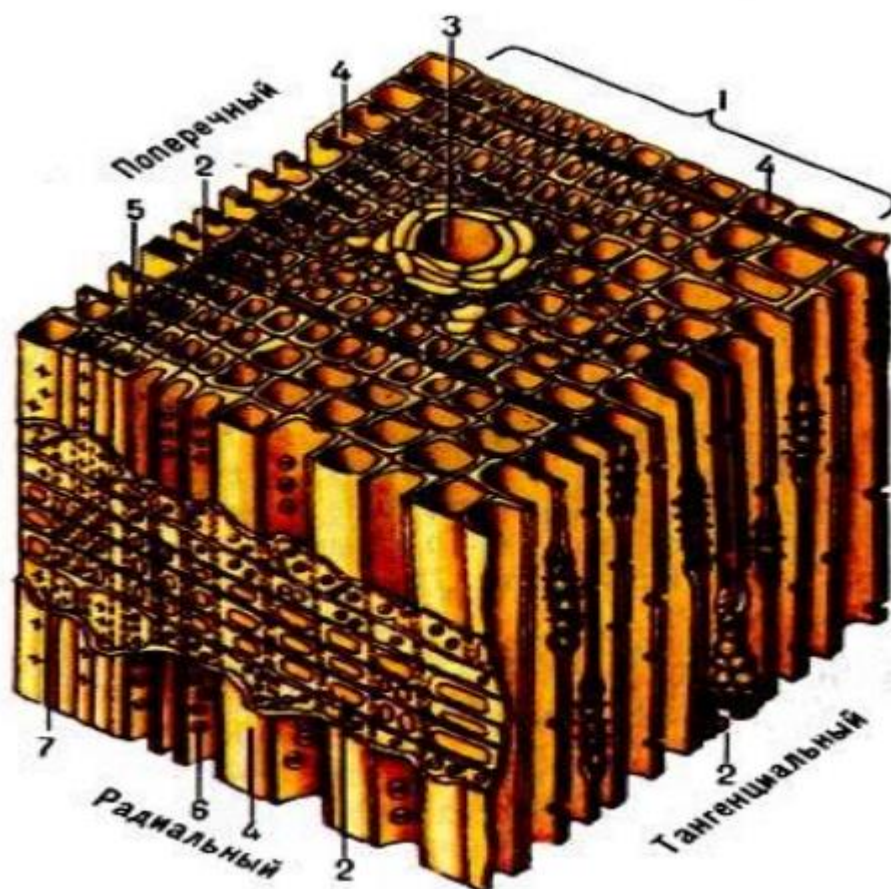


Рисунок 1 - Срез хвойной древесины

1 - годичный слой; 2 - многорядный луч с горизонтальным смоляным ходом; 3 - смоляной ход; 4 - ранние трахеиды; 5 - лучевые трахеиды; 6 - поры; 7 - сердцевинные лучи

Микроструктура древесины исследуется при помощи световой микроскопии на тонких срезах - поперечном, тангенциальном и радиальном, которые иллюстрирует рисунок 1. Микроскопические исследования показывают, что древесина состоит из большого количества плотно соединённых между собою клеток, отличающихся по размерам и форме.

Изм.	Колич.	Лист.	№ док	Подпись	Дата

ДП-230102.65-001019120ПЗ

Лист

7

90-95% объёма древесных клеток хвойных пород является мёртвыми клетками. Мёртвые клетки отличаются от живых прежде всего тем, что в их внутренних полостях отсутствует живая протоплазма и они заполнены водой или воздухом. Так как древесина в основном состоит из мёртвых клеток, то, очевидно, основная масса вещества древесины сосредоточена в клеточных оболочках.

На поперечном разрезе древесины вокруг сердцевины можно увидеть концентрические слои, называемые «годовыми кольцами» и представляющие собой ежегодный прирост древесины. Ширина годовых колец зависит от условий роста, положения в стволе, породы. Ширина годовых колец у одной и той же породы может различаться. При неблагоприятных условиях произрастания (нехватка питательных веществ, заболоченность почв, мороз, засуха) образуются узкие годовые слои.

1.1.1 Гистометрические измерения клеточных структур хвойных

90-95% объёма древесины хвойных составляют трахеиды, являющиеся её основными структурными элементами. Это вытянутые клетки со стенками различной толщины. Трахеиды подразделяются на ранние и поздние. Ранние трахеиды более широкие, их стенки тонкие, а полости широкие. Соотношение длины к ширине у таких клеток около 100). Поздние трахеиды заметно уже (соотношение длины к ширине приблизительно 200), они имеют толстые стенки и узкие полости. Поздние трахеиды выполняют механические функции. Размеры трахеид зависят от древесной породы и конкретных условий произрастания дерева.

Другим анатомическим элементом древесины хвойных являются паренхимные клетки, образующие живую ткань - паренхиму, клетки которой формируют сердцевинные лучи - ряды из паренхимных клеток, идущие горизонтально по радиусам ствола. У хвойных пород сердцевинные лучи узкие, однорядные, а их высота колеблется по числу клеток.

В древесине ряда хвойных пород (сосна, лиственница, ель) присутствует помимо прочего эпителиальная паренхима, образующая смоляные каналы. Это межклеточные каналы, заполненные смолой (живицей). Различают вертикальные и горизонтальные смоляные каналы, формирующие единую смолоносную систему. Вертикальные каналы располагаются чаще в поздней зоне годового кольца и образуются тремя слоями клеток: внутренним выстилающим слоем живых эпителиальных клеток; слоем мёртвых клеток, заполненных воздухом; слоем живых клеток, отмирающих со временем. Горизонтальные смоляные каналы проходят в многорядных сердцевинных лучах. Они образованы живыми эпителиальными клетками. Смолоносная система служит древесному растению защитной системой, позволяющей переносить экстремальные условия обитания, и в том числе обеспечивает защиту от повреждения дерева насекомыми и другими живыми организмами.

Клетки древесины связаны между собой через поры - неутолщённые участки клеточной стенки. Пора не является свободным отверстием, так как в ней имеется твёрдая мембрана (первичная стенка и межклеточное вещество), пронизанная микроскопическими отверстиями. В живых клетках через эти отверстия проходят тонкие нити цитоплазмы, соединяющие содержимое живых клеток в единое целое.

В условиях тёплого климата, где нет смены времён зима-лето, рост древесины происходит круглогодично. В холодной климатической зоне камбий функционирует периодически, то есть деятельность его замирает на зиму и возобновляется весной. Благодаря годичной периодичности работы камбия отлагаемая древесина образует так называемую годичную слоистость. В каждом годичном слое можно различить тонкостенную - раннюю (весеннюю) и толстостенную - позднюю (летнюю или осеннюю) древесину. Ранняя древесина, образуемая весной и в начале лета, занимает внутреннюю часть годичного слоя, а поздняя - наружную, ближайшую к камбию часть кольца.

Вариативность строения годичных слоёв деревьев хвойных пород под воздействием физических факторов внешней среды позволяет использовать их в качестве косвенных индикаторов изменения климата и условий среды в природных экосистемах на длительных временных интервалах. Использование годичных слоёв в качестве косвенных индикаторов эффективно по многим причинам. Они чувствительны к внешним воздействиям, имеют постоянство реакции на изменение климатических параметров среды. К достоинствам относится также возможность реконструкции события с точностью до года при использовании ширины годичного слоя, а при использовании структуры годичных слоёв появляется возможность реконструкции в границах вегетационного периода. Не менее важны относительная лёгкость взятия пробы и, как правило, возможность взять достаточное количество образцов для получения статистически достоверных измерений.

Ширина годичного слоя является первым, наиболее легко измеряемым и обладающим высокой чувствительностью параметром годичных колец. Ширина годичного слоя прямо пропорциональна количеству клеток в годичном слое и отражает интегральное воздействие факторов внешней среды на дерево в течение вегетационного периода. Существует очень значительное количество работ, описывающих зависимость ширины годичного слоя от факторов внешней среды, так как дендрохронология использует этот параметр с конца XIX - начала XX веков. В сущности, любое внешнее воздействие на древесное растение приводит к изменению ширины годичного слоя. Это могут быть температурный и гидрологический режимы зоны произрастания, промышленные выбросы и лесохозяйственные мероприятия, сход снежных лавин и камнепады, вспышки массового размножения вредителей и пожары, извержение вулканов и падение метеоритов, ультрафиолетовое и проникающее ионизирующее излучения, радиоактивное и химическое

загрязнения местности и т.д. Ухудшение условий роста приводит к снижению частоты делений боковой меристемы, что вызывает уменьшение ширины годичного слоя. В сезоны сильных депрессий роста годичный слой может быть представлен единственной клеткой или отсутствовать вовсе на данном радиусе ствола дерева. Наличие общего фактора или группы факторов, которые действуют на деревья, произрастающие в одном районе, вызывает синхронное изменение у них ширины годичного слоя, что позволяет получать усреднённые для данного участка хронологии ширины годичного слоя с незначительным содержанием шумов и максимальным выделением полезной составляющей.

В отличие от ширины годичного слоя, радиальный размер трахеид и толщина их клеточных стенок позволяют фиксировать влияние внешних факторов в течение вегетационного периода. Характерным примером этого служат так называемые «ложные кольца», образуемые мелкими толстостенными клетками в начале зоны поздней древесины и более крупными, сформированными позже. Описанный эффект наблюдается в условиях недостатка увлажнённости и отчётливо выраженного сезона дождей - уменьшение радиальных размеров трахеид вызвано сезонной засухой. Недостаток увлажнённости сказывается также и на распределении радиальных клеточных размеров в годичных слоях деревьев. Для годичных клеток сосны из проточных условий кривая распределения размеров клеток приближается к нормальному распределению со слабой левосторонней асимметрией, в то время как для умеренно увлажнённых мест возникает хорошо выраженная двухвершинность распределения, ещё более контрастная для сухого местообитания.

В случае, если в конце вегетационного периода наступает резкое похолодание, наблюдается формирование клеток с тонкими клеточными стенками у деревьев, растущих в высоких широтах или на верхнем горном пределе. Такие слои носят название «светлые кольца», так как они хорошо визуально определяются за счёт неразвитых клеток поздней древесины, поэтому такое кольцо по сравнению с остальными выглядит более светлым.

Чувствительность характеристик клеточной структуры к изменению физических факторов внешней среды стала основой имитационной модели роста дерева, рассчитывающей основные структурные параметры годичных слоёв по изменениям интенсивности влагообеспечения, температуры воздуха и солнечной радиации. Это впервые дало возможность проводить численное моделирование реакции деревьев на изменение параметров окружающей среды.

1.1.2 Другие методы измерения клеточных структур хвойных

Помимо гистометрических параметров годичных слоёв, со второй половины XX века начали разрабатываться новые физические методы и приборы для исследования влияния факторов среды на рост деревьев.

Наиболее распространённое применение на практике получил метод рентгенографии, в основе которого лежит денситометрия с использованием негативов изображений древесных образцов, полученных в рентгеновском диапазоне длин волн на рентгеновских плёнках. В результате применения методов рентгеновской денситометрии было показано, что изменение плотности древесины внутри годовичных слоёв часто является лучшим показателем условий среды, чем ширина годовичного слоя. Также можно отметить способ измерения плотности древесины через ослабление потока электронов, метод профилометрии образца древесины для оценки размеров элементов и т.п.

Близкими рентгенографическому методу исследования годовичных слоёв можно считать микрофотометрический метод измерения оптической плотности древесных срезов в проходящем свете и разработанный в Красноярске микрофотометрический метод анализа структуры годовичных слоёв в отражённом свете. Указанные методы дают возможность получать профили яркости световых потоков годовичных колец, проходящих через образец или отражённых от его поверхности. Обратная величина интенсивности этих колец имеет большое сходство с профилем плотности годовичных слоёв, полученных с помощью рентгеновской денситометрии. Проведённые исследования показали, что возможно применение профилей яркости годовичных колец хвойных в исследовании строения и свойств древесины, анализе сезонного роста дерева, исследовании влияния на сезонный рост дерева погодных условий, насекомых-фитофагов и других факторов. Широкому применению данных методов помешали трудности, связанные с приготовлением образцов. Среди них можно указать на невозможность приготовления длинных микротомных препаратов, неравномерность окраски лучей и других клеточных элементов в годовичном слое, различие отражательной способности участков древесины в разных частях дерева (например, в заболони и в ядре).

1.2 Применяемая методика измерения структур клеток хвойных

В монографии Силкина П. П. «Методы многопараметрического анализа структуры годовичных колец хвойных» выделены следующие основные измеряемые структурные элементы трахеид, которые ниже иллюстрирует рисунок 2:

- Lu - радиальный размер люмена;
- D - радиальный клеточный размер;
- T - тангенциальный клеточный размер;
- W - толщина стенки клетки;
- WW - толщина двойной стенки клетки.

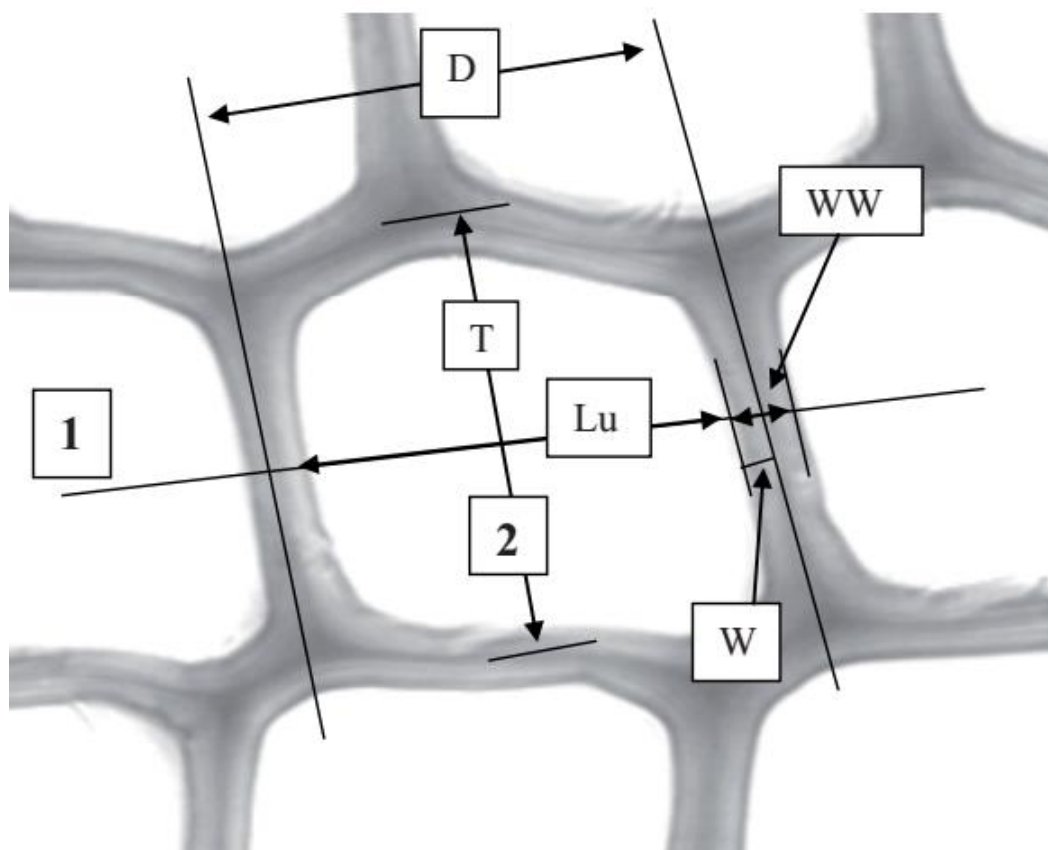


Рисунок 2 - Измеряемые структурные элементы трахеид хвойных

Для того чтобы результаты измерений можно было корректно использовать, в данной работе также устанавливаются следующие правила измерений клеточных структур:

- измерение радиального клеточного размера (D) и радиального размера люмена (Lu) проводятся вдоль радиальной линии, перпендикулярной стенкам клеток. Точки пересечения линии с стенками клеток должны проходить в средней части стенок клеток;
- измерение толщины стенки клетки (W) проводится в средней части стенок клеток по радиальной линии. Основное условие - перпендикулярность линии измерения стенкам клеток;
- измерение тангенциального клеточного размера (T) проводится вдоль линии, перпендикулярной радиальной линии и делящей её пополам.

В разных рядах одного годичного слоя присутствует естественный разброс клеточных размеров и толщины стенки клетки, поэтому требуется измерить несколько рядов клеток для расчёта средних величин клеточных размеров для каждого номера клетки в радиальном ряду. Данной проблеме исследователи уже уделяли внимание. Например, в указанной выше работе Силкина П. П. указано, что средняя выборочная трахеидограмма по 6 рядам клеток в годичных слоях сосны отличается от средней генеральной не более чем на 5%. Тем не менее важно сделать оценки необходимого количества измерений рядов клеток для разных пород деревьев.

Исследования в указанной выше работе Силкина П. П. проводились для двух пород хвойных - лиственницы и сосны. Были измерены радиальные размеры клетки и толщина клеточной стенки для 30 рядов в каждом годичном слое. Всего было измерено 20 случайно выбранных годичных слоёв у 5 деревьев обеих пород. Для всех годичных слоёв была определена относительная погрешность измерения при разном числе рядов. Полученные значения группировались по участкам годичного слоя, отдельно для каждой породы. Для каждой группы при каждом фиксированном количестве рядов клеток были рассчитаны стандартные статистические характеристики согласно формулам статистики. Результаты расчётов показали, что для всех участков годичного слоя у лиственницы и сосны в большинстве случаев можно ограничиться измерением 5-6 рядов клеток, так как дальнейшее увеличение количества измеряемых рядов не приводит к существенному увеличению точности измерения.

Другой измеряемой характеристикой годичных слоёв является тангенциальный размер клеток. Установлено, что средний тангенциальный размер клеток остаётся неизменным как в рамках годичного слоя, так и на всём радиусе ствола дерева. Существуют небольшие вариации среднего тангенциального размера между отдельными деревьями. В монографии Силкина П. П. «Методы многопараметрического анализа структуры годичных колец хвойных» приведена оценка необходимого числа измерений тангенциального размера у радиальных рядов клеток в годичных слоях лиственницы и сосны. Отдельно измерялись тангенциальные размеры у зоны ранней и поздней древесины. Слои выбирались случайным образом по краям образца и в его середине. Оценки показали, что для измерения среднего тангенциального клеточного размера с надёжностью 95% и относительной погрешностью 5% достаточно измерить 120 рядов в годичных слоях лиственницы и 103 ряда в годичных кольцах сосны.

Изм.	Колич.	Лист.	№ док	Подпись	Дата

ДП-230102.65-001019120ПЗ

Лист

13

2 Обзор существующих решений в области анализа клеток

Количественная фотоморфометрия изображений в поле зрения микроскопа является одним из основных современных методов работы с цитологическими препаратами в биологических исследованиях и медицинской диагностике. При этом класс задач, которые могут быть решены при помощи этих методов, достаточно широк и включает в себя такие важные проблемы, как определение раковых клеток, распознавание клеток крови и костного мозга, выявление нарушений тканей и так далее. В последние тридцать лет предприняты многочисленные попытки автоматизировать этот процесс с целью применения в медицинской диагностике широкого круга заболеваний.

В настоящее время в ряде ведущих отечественных (в меньшей степени) и зарубежных центров ведутся крупномасштабные исследования по разработке приборов, автоматизирующих распознавание и измерение микроскопических изображений клеток в цитологических и гистологических препаратах. Использование вычислительной техники и математических методов в этой отрасли позволяет автоматизировать этот процесс, повысить точность результатов исследований и провести соотнесение выявленных особенностей изображений с результатами прочих методов исследования (физиологических, биофизических, биохимических, и т.д.), сделать результаты исследований более объективными. Помимо этого, загрузка изображений в компьютер не только позволяет сортировать и хранить их, но и осуществлять обмен другими исследователями при помощи современных средств компьютерных телекоммуникаций. Существующий опыт доказывает, что разработка автоматического цитофотоморфометра - весьма сложная задача с математической и технической точек зрения. Это объясняется высокой вариабельностью клеток и клеточных структур, высоким уровнем шумов и помех, частичной утратой информации в фотоизображениях по сравнению с трёхмерным оригиналом, неравномерностью условий окраски в процессе приготовления образцов и т.п. На сегодняшний день задача разработки подобной системы решена лишь частично, прежде всего из-за сложности автоматического распознавания и измерения слабоконтрастных цитологических образов. Как раз этим фактом можно объяснить безраздельное господство в медицинской практике высокопроизводительных проточных цитоанализаторов, несмотря на такие очевидные плюсы цитоанализатора изображений как гибкость и количество выделяемых клеточных характеристик.

Одним из наиболее распространённых и мощных инструментов биологических исследований является оптическая микроскопия. Начиная с 70-х годов прошлого века, предпринимались попытки компьютерного анализа изображений поля зрения микроскопа. Получивший распространение термин «автоматизированная микроскопия» («automated microscopy») соответствует системе микроскопии, в которой изображение поля зрения преобразуется в

цифровую форму, позволяющую использовать компьютер для автоматизации микроскопического анализа и управления микроскопом. Однако создание полностью автоматизированных систем микроскопии стало возможно лишь в последнее десятилетие благодаря бурному прогрессу в области вычислительных средств, техники видеосъёмки, математических методов анализа.

Одна из главных задач построения систем автоматизированной микроскопии - выделение объектов на изображениях. Так, например, выделение гистологических объектов на цифровых изображениях затруднено их существенной вариабельностью и слабой контрастностью. Кроме того, границы самих микрообъектов и границы их внутренних элементов изначально, по своей природе, нечёткие. В настоящее время предложено большое число подходов к решению этой проблемы. Предлагаются методы сегментации с использованием различных фильтров-детекторов границ, сегментация методом k-средних, использование искусственных нейронных сетей, метод роста регионов и другие. Универсальные подходы к выбору алгоритма для сегментации произвольного изображения гистологических объектов неизвестны.

Для построения вектора признаков, используемого для классификации обнаруженных объектов, применяют как морфологические признаки, так и признаки, основанные на текстурных и цветовых особенностях. В качестве классификаторов используются разного рода нейронные сети.

Некоторые из упомянутых методов получили реализацию в так называемых автономных системах автоматизированной микроскопии (АСАМ). АСАМ автоматически перемещает и фокусирует препарат, в зависимости от распределения биоматериала выбирает траекторию просмотра, контролирует качество окраски и освещения, распознаёт и сохраняет в базу данных объекты заданных типов. Выполняется измерение и анализ автоматически собранной выборки объектов. Кроме автоматизированного микроскопа и системы анализа изображений в состав АСАМ входит программный компонент, управляющий роботизированным микроскопом и заменяющий при перемещении препарата глаза и руки лаборанта. Примерами таких систем являются:

- «DiffMaster Octavia» (Швеция) для автоматического подсчёта лейкоцитарной формулы по мазку крови. В системе используется система классификации на основе обучающихся нейронных сетей;
- «Rare Event Imaging System» (REIS, США) для автоматического обнаружения инфицированных цитомегаловирусом лейкоцитов периферической крови. Морфометрические признаки помеченных лейкоцитов используются в автоматической диагностической системе;
- «IQ-200» (IRIS, США) для автоматического анализа осадка мочи. Система обнаруживает 12 элементов мочи - эритроциты, лейкоциты, лейкоцитарные сгустки, бактерии, грибки, гиалиновые цилиндры, другие

патологические цилиндры, клетки плоского эпителия, другие эпителиальные клетки, слизь, сперматозоиды и артефакты;

- «МЕКОС-Ц1» (МЕКОС, Россия) для автоматического подсчёта лейкоцитарной формулы и количественного анализа эритроцитов и ретикулоцитов. В системе используется система классификации на основе обучающихся нейронных сетей.

Сегодня существует большое количество методов сегментации цитологических изображений, выбор которых определяется особенностями поставленной задачи. К примеру, для контрастно окрашенных препаратов часто применяются хорошо зарекомендовавшие себя методы пороговой обработки вкупе с морфологическими операциями эрозии и дилатации. В этих методах в качестве параметров используются распределения оптических плотностей объектов. Эти методы эффективны в ситуациях, когда распределения оптической плотности различных сегментируемых объектов имеют малозначительные пересечения. Преимуществом подобных методов является теоретическая проработка и высокая скорость обработки изображений. Такой метод достаточно эффективен для сегментации сравнительно простых объектов, таких как изолированные эритроциты или изолированные ядра. Однако в более сложных случаях, к примеру, при исследовании параметров ядрышек и ядер клеток белой крови, даже в случае отдельно лежащих клеток практически при окраске любого типа некоторые объекты имеют слабоконтрастные границы и значительные пересечения распределений оптических плотностей различных объектов (цветностей в случае цветных изображений). В этих случаях эффективность указанных методов невелика. Для подобных задач приходится применять более серьёзные алгоритмы, использующие помимо параметров распределения оптических плотностей геометрические свойства границ сегментируемых объектов. Примером может служить работа по сегментации изображений в маммологии. Для цитологических исследований, когда необходимо обработать большое количество изображений, такие методы в практической реализации до сих пор не применялись. Это связано, судя по всему, со сложностью существующих общих методов и отсутствием для них эффективных вычислительных систем.

2.1 Решения в области анализа клеточных структур крови

В настоящее время в медицинской практике, помимо «классического» метода микроскопического подсчёта клеток в специально окрашенном мазке взятой из пальца крови, используют две группы методов автоматизированного анализа и подсчёта клеток крови: проточную цитометрию и компьютерный анализ изображения. Они позволяют унифицировать упомянутый «классический» метод визуального анализа, что значительно увеличивает точность получаемых результатов.

Изм.	Колич.	Лист.	№ док	Подпись	Дата

ДП-230102.65-001019120ПЗ

Лист

16

2.1.1 Проточная цитометрия

Метод проточной цитометрии основывается на флуоресцентной колоризации клеточной РНК с последующим обнаружением свечения. Его использование позволяет не только определить количество клеток, но и дать оценку степени их зрелости по количеству содержащейся в клетках РНК.

Современные методы проточной цитометрии начали своё развитие с 1956 года, когда был создан прибор для автоматического подсчёта и анализа кровяных клеток. Сам принцип проточной сортировки клеток, анализируемых в потоке, разбитом на мелкие капли, был реализован в 1965 г. Эти капли несли на своей поверхности заряд, позволяющий производить их сортировку по разным пробиркам. Принцип проточной цитометрии заключается в быстром пропуске клеточной суспензии через область чувствительности прибора. Концентрация клеток подбирается таким образом, чтобы скорость регистрации была достаточной, но такой, чтобы не допустить при этом появления в области обнаружения одновременно двух объектов. Данные от датчиков в импульсной форме фиксируются счётчиками, анализаторами импульсов или ЭВМ (с оцифровыванием). Информация накапливается и представляется в виде гистограммы частот распределения.

Сегодня системы проточной цитометрии достаточно широко используются в биологических и медицинских исследованиях. Развитие технического оснащения и методов маркирования клеток открыли уникальные возможности для анализа огромного количества различных клеточных характеристик, что облегчает получение новых знаний о процессах, обеспечивающих функционирование клеток и клеточных структур. В гематологии проточные системы используются для быстрого автоматического счёта элементов крови, дифференциального анализа (подсчёт формулы крови, выделение атипичных клеток и их фенотипирование), изучения клеточного цикла и многого другого.

Аппаратура для проточной цитометрии постоянно совершенствуется в направлении конструирования более удобных для применения специализированных автоматизированных систем, позволяющих повысить точность и скорость обработки и не требующих высокой квалификации обслуживающего персонала. Высокий технический уровень оснащения для проточной цитометрии уже достигнут, и дальнейшее развитие этого направления идёт по пути разработки новых методов флуоресцентной маркировки биологических объектов, создания более узкоспециальных и чувствительных зондов, усовершенствования методик фиксации и окраски клеток.

На сегодняшний день большинство приборов для проточной цитометрии выпускаются двумя зарубежными производителями: Becton Dickinson и Coulter. В России проточные цитофлуориметры производятся предприятием «Медимакс».

Изм.	Колич.	Лист.	№ док	Подпись	Дата

ДП-230102.65-001019120ПЗ

Лист

17

Помимо высокой стоимости аппаратуры, недостатком систем проточной цитометрии можно считать отсутствие прямой визуализации исследуемых объектов. Однако этот недостаток нивелируется большими объёмами изучаемых клеток. Кроме того, представляющие интерес объекты, могут быть выделены и изучены другими методами, в том числе при помощи компьютерного гематологического анализа.

2.1.2 Компьютерный гематологический анализ

В медицинской практике значительное распространение получили приборы для быстрого автоматического подсчёта и анализа элементов крови - гематологические анализаторы. В своём большинстве это сравнительно дешёвые и несложные в эксплуатации приборы, позволяющие проводить рутинные исследования, связанные с подсчётом количества и определением размеров клеток крови. Данные приборы выпускаются производителями, специализирующимися на лабораторном медицинском оборудовании: «Abbott», «Analyse Instrument AB», «Coulter», «Medicor», «Roch», «Seron», «Sysmex» и другими.

Гематологические анализаторы - это специализированные автоматизированные системы с компьютерной обработкой сигналов, осуществляющие многопараметрическую оценку данных в сочетании с визуальным представлением основных клеточных популяций. Характерными особенностями этой аппаратуры являются:

- автоматическое взятие и развитие пробы крови;
- применение методов предварительной обработки пробы;
- использование нескольких каналов для обнаружения сигналов (поляризация светорассеивания, радиочастотный анализ, кондуктометрическое измерение объёма, учёт поглощения и рассеивания света и т.п.);
- применение специальных вычислительных программных алгоритмов для подсчёта клеток и предоставления окончательных результатов.

Например, гематологический анализатор «Cobas Vega» (Roche, Швейцария) даёт возможность анализировать 22 гематологических характеристики, а также вычислять абсолютное и процентное содержание 2-х патологических клеточных популяций: атипичных лимфоцитов и больших незрелых клеток.

К анализаторам автоматического (компьютеризованного) типа принадлежит отечественная система «АСПЕК», разработанная и созданная ОАО «Радиотехнический институт имени академика А.Л. Минца» совместно с ФГБУ «Гематологический научный центр» Министерства здравоохранения РФ, способная анализировать по мазку, взятому из некоторого объема пробы периферической крови, её клеточный состав. В режиме автоматической микроскопии прибор осуществляет: сканирование препарата по заданному

маршруту; автоматическое фокусирование; ввод серии кадров микроскопа видеокамерой; обработку этих изображений на ЭВМ. Последний этап основан на следующей схеме: входное изображение сегментируется, на нём выделяются связанные зоны, среди которых посредством алгоритмов распознавания идентифицируются клетки крови.

Сегментация даёт возможность выделить базовые элементы кадра; идентифицировать участки, соответствующие клеткам, фону, артефактам; учитывать внутриклеточные структуры для полного описания объекта при подсчёте лейкоцитарной формулы.

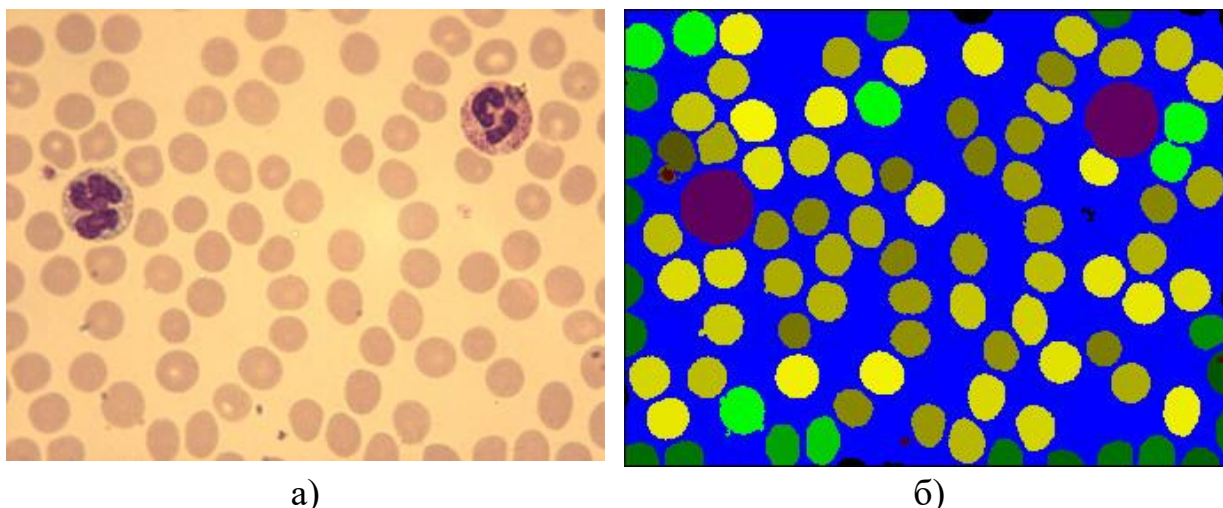


Рисунок 3 - Пример работы алгоритма сегментации комплекса «АСПЕК»
а - изображение поля зрения микроскопа; б - результат сегментации

Алгоритм сегментации не требует информации о правилах распределения и количестве сегментируемых участков. Поэтому он устойчив к естественной вариативности объектов, изменению насыщенности и условий окрашивания препаратов, а также условий съёмки изображений. В установке применена модель равномерно распределённых, независимых гауссовских точек. Это сделано для повышения скорости и уменьшения сложности вычислений; но, тем не менее, для морфометрического анализа применяются и более сложные модели анализа изображений. Подсчёт форменных элементов производится покадрово; и если изображение содержит области, идентифицируемые как лейкоциты, то эти области сегментируются снова для обнаружения внутриклеточных структур. После чего обнаруженные внутриклеточные структуры распознаются как ядра и области цитоплазмы, которые затем измеряются. Опираясь на полученные данные посредством пошаговой процедуры распознавания вычисляется лейкоцитарная формула.

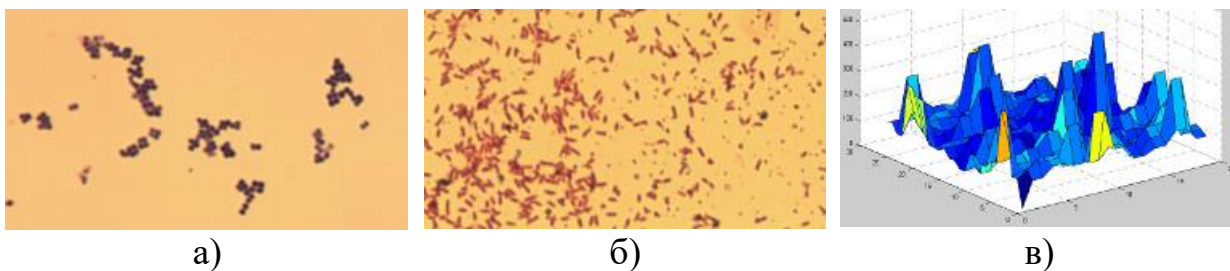


Рисунок 4 - Результаты работы алгоритмов распознавания комплекса «АСПЕК» (а, б - изображения микробных объектов; в - распределение по мазку)

Методы распознавания в установке «АСПЕК» должны обеспечивать быстроедействие. Для этой цели применяется двухэтапный алгоритм:

- первый этап - построение простого и быстрого алгоритма распознавания, основанного на априорных данных о кровяных клетках в мазке: их идентификация (эритроциты, тромбоциты, лейкоциты) и подсчёт их количества в кадре и на всей площади мазка;
- второй этап - вычисление лейкоцитарной формулы, при котором используется традиционная статистическая многоэтапная процедура классификации. Описание объекта при этой процедуре значительно полнее, количество классов значительно больше и может наращиваться в процессе применения прибора.

В целях оптимизации процедуры классификации априорные данные разделяют на несколько шагов и каждый шаг анализируют отдельно. Описанная схема имеет методическое сходство с «древовидными» процедурами, подразумевающими увеличение количества распознаваемых классов, и, вместе с тем, характеризуется быстроедействием и стойкостью к вариативности объектов. На каждом шаге классификации изображений лейкоцитов они группируются по наборам свойств в промежуточные классы. Классифицированные подобным образом клетки передаются на следующие этапы, где их классификация уточняется с применением следующей группы свойств. В качестве примеров промежуточных классов можно привести изображения многоядерных клеток, имеющих ядро правильной формы и другие особенности. Достоинствами описанного алгоритма являются минимизация количества используемых свойств без ущерба качеству анализа; возможность видоизменения процедуры посредством изменения количества этапов и наборов цитологических свойств; простая программная реализация процедуры и возможность её видоизменения без дополнительного программирования.

Разумеется, анализатор изображения «АСПЕК» проигрывает проточным счётчикам в скорости, но при этом имеет ряд преимуществ: он может распознавать большое количество видов клеток; он «обучаем»: в процессе работы расширяется количество распознаваемых разновидностей клеток. Крайне важную роль имеют визуализация результатов исследования и возможность корректировки результатов; компьютеризация сильно расширяет

базу получаемой при исследовании информации, при этом остаётся возможность сравнения результатов; применяемые материалы относительно недороги.

2.2 Решения в области распознавания и анализа клеточных структур деревьев хвойных пород

Типичный размер клеток в годовичных слоях деревьев хвойных пород колеблется в пределах 10-50 мкм, а толщина клеточной стенки - в пределах 2-10 мкм, поэтому измерение микроструктуры проводится с использованием световой микроскопии. В исследовательских лабораториях используют методы световой микроскопии как проходящего, так и отражённого света. И тот и другой метод имеет свои достоинства и недостатки. В случае исследования годовичных колец в проходящем свете готовится препарат по стандартной методике. Препарат представляет собой окрашенный тонкий, порядка 20-25 мкм, поперечный срез древесины, помещённый между предметным и покровным стёклами в какую-либо оптически прозрачную среду. Этим достигается хорошая контрастность изображения при которой можно наблюдать структуру клеточных стенок, а также соблюдается постоянство расстояния между объективом и препаратом.

Современное развитие микроэлектроники и компьютерных технологий привели к созданию автоматизированных комплексов, в состав которых входят световой микроскоп, оснащённый видеокамерой, компьютер и специализированное ПО, позволяющее управлять режимами работы микроскопа, считывать и обрабатывать изображение, сформированное видеокамерой. Общепринятым термином для названия описанных комплексов в нашей стране является «система анализа изображений» (САИ). За рубежом используется аналогичные термины «Image-System» или «Imaging System». В лаборатории функционирования лесных экосистем Института фундаментальной биологии и биотехнологии Сибирского федерального университета также существует САИ. Рисунок 3 показывает внешний вид комплекса, а рисунок 4 - его структурную схему. В состав данного комплекса входят компьютер, световой микроскоп проходящего и отражённого света «AxioImager.D1», оснащённый цветной видеокамерой «AxioCam MRc5», универсальная компьютерная программа «AxioVision» (Карл Цейс, Германия), а также специализированный пакет прикладных программ. Необходимость в разработке собственного специализированного пакета программ диктовалась тем, что любые универсальные программы, решающие задачи широкого круга специалистов, как правило, малоэффективны при проведении массовых измерений в определённых задачах, к коим относится измерение клеточной структуры годовичных слоёв хвойных. Разработанный пакет программ является аппаратно-независимым и легко адаптируется к любой базовой комплектации САИ.

Чтобы было легко проводить измерения клеточных структур, требуется получать как можно более контрастные, чёткие изображения годовичных слоёв без посторонних деталей. Идеальный вариант соответствовал бы ситуации, когда люмен был бы представлен абсолютно белым изображением, а клеточная стенка - абсолютно черным (бинарное изображение). Для повышения контрастности изображения чаще всего используются световые фильтры, что, безусловно, повышает контраст между стенкой клетки и люменом. К сожалению, часто этого бывает недостаточно.

Для преобразования изображений годовичных колец в бинарные, обычно используют спектральную обработку этих изображений на компьютере. В реальности такая методика применима только для безупречно сделанных изображений годовичных колец, а значит, она неприемлема для проведения измерения больших массивов клеток. Например, небольшую тень у клеточной стенки алгоритмы ПО превратят в элемент клеточной стенки, что приведёт к значительным погрешностям измерения её толщины.

Для получения изображений с требуемым качеством обычно применяются два доступных способа увеличения контрастности изображения:

- цветные светофильтры.
- увеличение светового потока от источника света до такого уровня, когда световой поток, проходящий через люмен, превышает верхний рабочий порог матрицы видеокамеры, что приводит к её «ослеплению». При таком способе люмен выглядит равномерно-белым, без посторонних деталей.



Рисунок 5 - Комплекс получения и обработки изображений
микроскопических объектов (САИ)

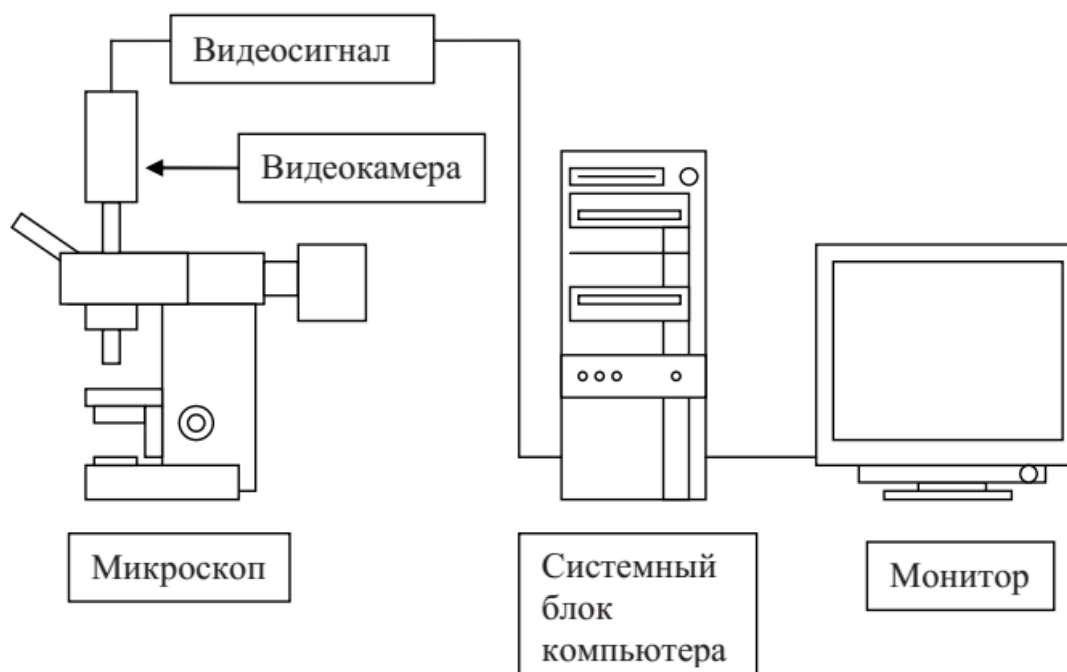


Рисунок 6 - Структурная схема комплекса

Изм.	Колич.	Лист.	№ док	Подпись	Дата

ДП-230102.65-001019120ПЗ

Лист

23

Благодаря подобному подходу отпадает необходимость в использовании спектральной обработки изображения, так как оно получается достаточно высокого качества и по своему виду приближается к бинарному. Реальное использование описанной методики получения фотоизображений годичных колец хвойных, описанное в монографии Силкина П. П. «Методы многопараметрического анализа структуры годичных колец хвойных» показало её высокую эффективность. Для последующей обработки полученных фотоизображений использовался пакет специализированного ПО, состоящий из программ «SuperMoment», «Lineyka», «ProcessorKR». «ProcessorKR» производит обработку измеренных величин (вычисление ширины годичного кольца, датировку клеточных данных, стандартизацию клеточных размеров, усреднение в пределах кольца и т.п.).

2.2.1 Программа «SuperMoment»

Одна из проблем при работе с годичными кольцами - невозможность охватить одним снимком с приемлемым разрешением всю интересующую область годичного кольца. Как правило, существует необходимость сделать несколько кадров. Для совмещения разрозненных изображений годичного кольца в одно было разработана программа «SuperMoment», предназначенная для автоматической компиляции отдельных изображений годичных колец в одно целое. Эта программа работает в паре со стандартными, базовыми программами САИ. Оператор с помощью базовой программы создаёт цепочку файлов с изображением участков годичного кольца на жёстком диске компьютера. «SuperMoment» считывает изображения с диска, определяет их смещения друг относительно друга по двум, взаимно ортогональным осям и сохраняет изображения в виде одного файла, с учётом их взаимного расположения. Если одной линии фотографирования годичного кольца бывает недостаточно, вся процедура повторяется для другой группы рядов клеток того же годичного кольца. Рисунок 7 иллюстрирует пример сформированного изображения годичного кольца.

Программа «SuperMoment» имеет удобный графический интерфейс и позволяет оператору быстро проводить фотографирование годичных колец. В случае затруднений есть возможность ручного редактирования скомпилированного изображения. Процесс автоматического анализа изображений и их склеивание занимают менее секунды для нескольких десятков кадров для любого компьютера. Точность склеивания кадров фактически абсолютна для подавляющего числа годичных колец.

2.2.2 Программа «Lineyka»

Программа «Lineyka» разработана для применения в исследованиях в области дендрохронологии и нацелена на измерение клеточных структур. Данная программа способна работать с любыми стандартными форматами

графических файлов, содержащих изображения годовых колец. Принцип измерения расстояния основан на подсчёте числа точек между концами измеряемого отрезка с последующим умножением полученного числа на коэффициент масштаба точки изображения. Произведение равно искомому расстоянию. Для определения размеров точки изображения служит процедура калибровки. Для этой цели используется изображение шкалы объект-микрометра в горизонтальном и вертикальном положении. В процессе калибровки измеряется количество точек между минимальными делениями шкалы, с расстоянием 10 мкм. Проводится многократное измерение различных участков калибровочной шкалы. Отношение измеряемого расстояния (10 мкм) к среднему количеству точек определяет размеры точки изображения. Процедура калибровки проводится отдельно для горизонтальной и вертикальной оси.

Рисунок 7 - Пример сформированного изображение годичного кольца сосны

Ручной режим измерения является вспомогательным и применяется в тех случаях, когда невозможно использовать полуавтоматический режим, например, при некачественном изображении годичного слоя.

того как линия трассировки задана, оператор запускает процедуру построения профиля «яркости» точек вдоль этой линии. В результате получается профиль, который демонстрирует рисунок 8. Данный профиль показывает уровень «серости» точек, так как высокие значения профиля соответствуют более тёмным участкам. В результате пикам профиля соответствуют клеточные стенки, а толщина пика - толщине стенки клетки.

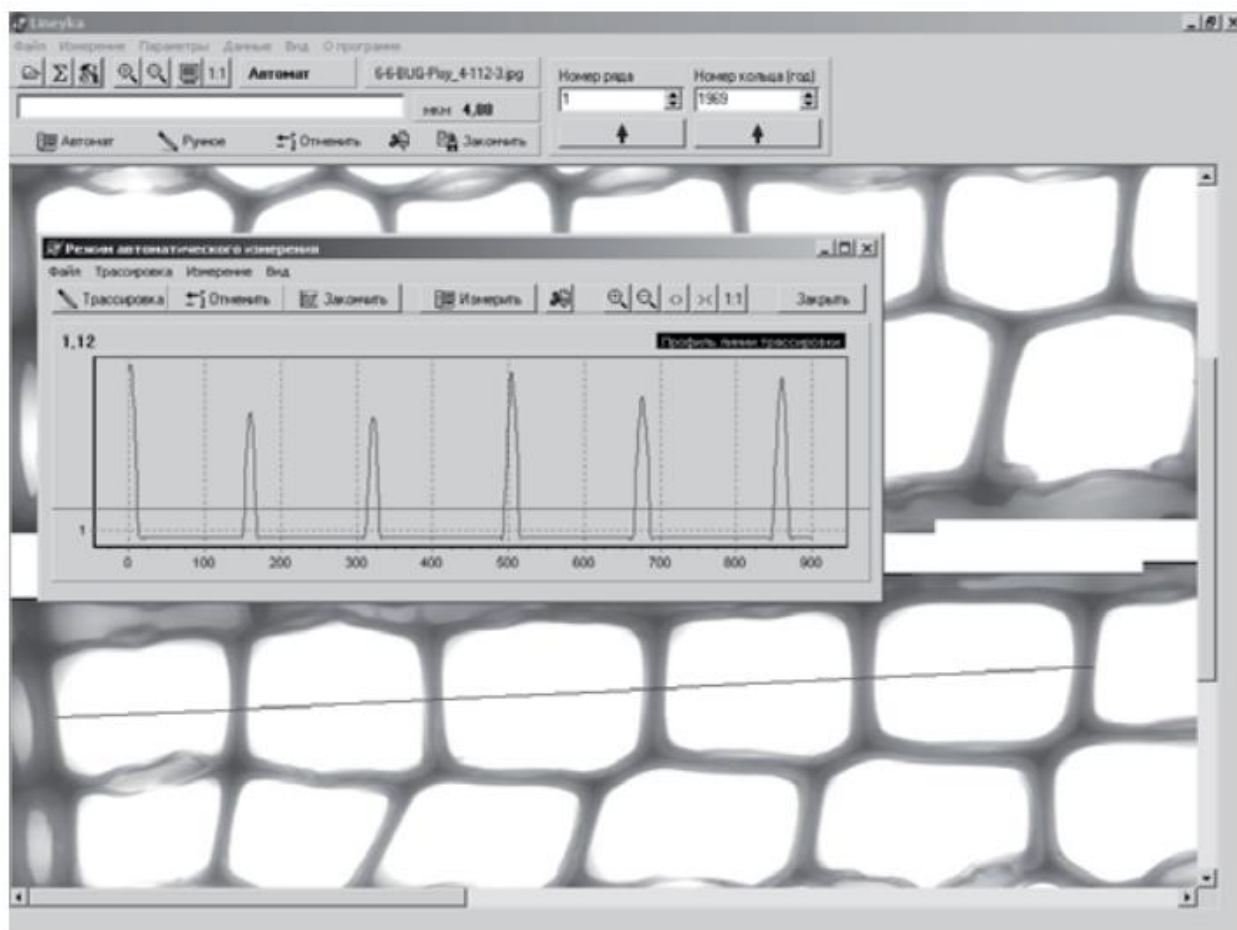


Рисунок 8 - Внешний вид программы «Lineyka» при работе в полуавтоматическом режиме

Затем оператор по профилю яркости задаёт линию, показывающую критический уровень яркости. Точки профиля, лежащие ниже этой линии, программа будет относить к люмену, а лежащие выше - к клеточной стенке. После включения процедуры измерения программа вычисляет размеры люменов и толщину клеточных стенок. Измеренные значения она отображает на самом рисунке в виде линий разного цвета. Например, рисунок 8 демонстрирует, как программа показала измеренные размеры люмена линиями синего цвета, а толщину клеточной стенки - красного, что позволяет оператору судить о качестве измерения и, в случае ошибки, подобрать другой критический уровень яркости и повторно провести измерения. В данном режиме измерение производится достаточно быстро и является гораздо менее утомительным способом измерения клеточных структур, в отличие от

полностью ручного режима. Наибольшую эффективность полуавтоматический режим измерения показал для образцов сосны, годовичные кольца которой содержат большие массивы клеток. Измеренные значения программа «Lineyka» записывает в текстовый тип файлов с расширением *.RealRec. Оператор имеет возможность ввести различные данные о препарате, такие как, к примеру, номер дерева, номер препарата, его тип, место сбора и т.д. Измеренные значения записываются поблочно для каждого годовичного кольца. Каждому измеренному значению ставится в соответствие номер ряда клеток, номер клетки в ряду, тип измеренной величины (люмен или стенка клетки).

						ДП-230102.65-001019120ПЗ	Лист
Изм.	Колич.	Лист.	№ док	Подпись	Дата		27

3 Анализ методов обработки и распознавания изображений

Одним из наиболее важных этапов автоматизированной обработки цитологических объектов является предварительная обработка изображений, которая позволяет повысить их качество, а также изменить некоторые характеристики таким образом, чтобы при количественном анализе изображений выделение интересующих объектов в дальнейшем могло бы быть произведено наилучшим образом. Для преобразования изображений используются различные методы фильтрации, которые можно условно разделить на четыре группы:

- изменение яркости и контрастности изображений. При использовании этих методов преобразования меняются яркостные характеристики изображения, что отражается на гистограмме яркости;
- фильтры выравнивания, обеспечивающие выравнивание изображений с точки зрения яркости и цвета, что необходимо в тех случаях, когда, например, имеются дефекты осветительной системы или оптики регистрирующей аппаратуры, выражающиеся в том, что фон изображений становится неравномерным;
- фильтры сглаживания, обеспечивающие очистку изображений от помех;
- фильтры детализации, обеспечивающие пограничные преобразования, что может выражаться (в зависимости от применённого фильтра) в усилении резкости, выделении границ объектов, детализации объектов и т.п.

3.1 Методы повышения контраста изображения

Наиболее распространённым недостатком фотографических изображений является низкий контраст, вызванный ограниченностью воспроизводимого яркостного диапазона. Под контрастом принято понимать разницу между максимумом и минимумом яркости. С помощью цифровой обработки контрастность можно усилить, меняя яркость каждой точки изображения и повышая яркостной диапазон. Для этого разработано несколько методов.

Допустим, к примеру, что уровни некоего изображения в оттенках серого занимают промежуток от 6 до 158 со средним значением яркости 67 при возможном полном диапазоне значений от 0 до 255. Рисунок 9а демонстрирует гистограмму яркостей исходного изображения, показывающую, какое количество точек N с близким значением яркости f укладывается в диапазон $[f_i, f+\Delta f_i]$. Данное изображение является слабоконтрастным, преобладают тёмные оттенки. Одним из способов повышения контрастности может стать так называемая «линейная растяжка гистограммы», когда уровням исходного изображения, лежащим в промежутке $[f_{\min}, f_{\max}]$, присваиваются новые значения с тем, чтобы охватить

весь полный диапазон изменения яркости, в данном случае от 0 до 255. При этом контрастность значительно усиливается. Модификация уровней яркости выполняется по формуле:

$$g_i = p + r f_i, \quad (1)$$

где

f_i - исходная величина яркости i -го пиксела;

g_i - модифицированная величина яркости;

p, r - коэффициенты.

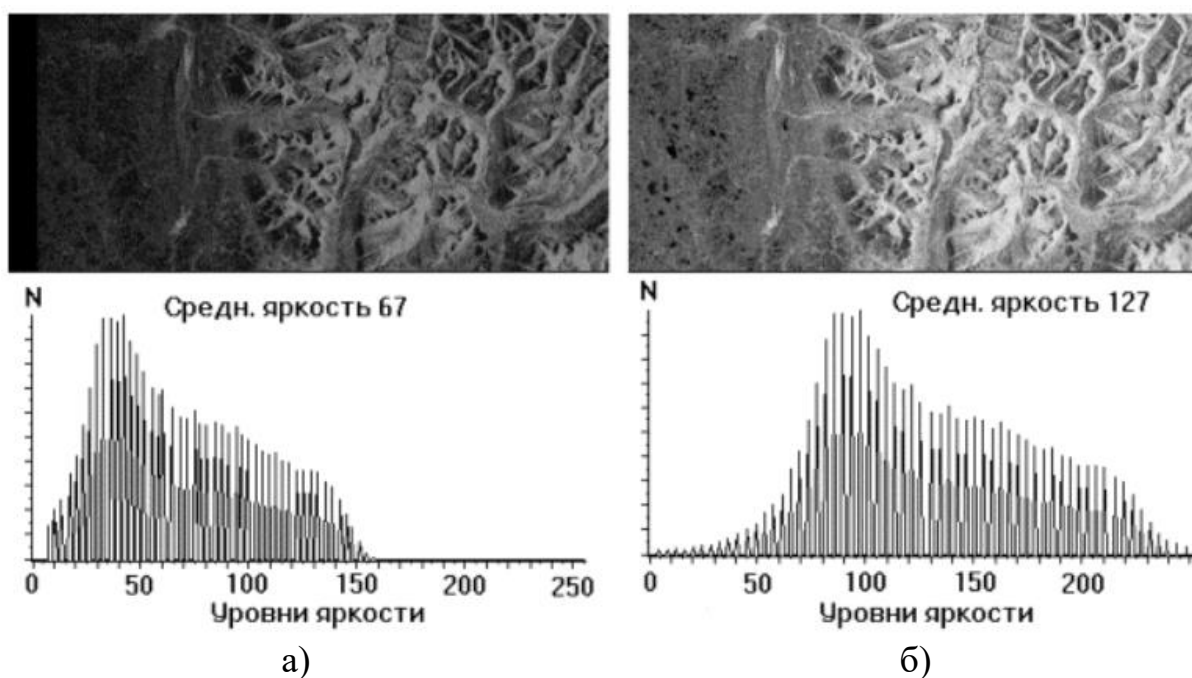
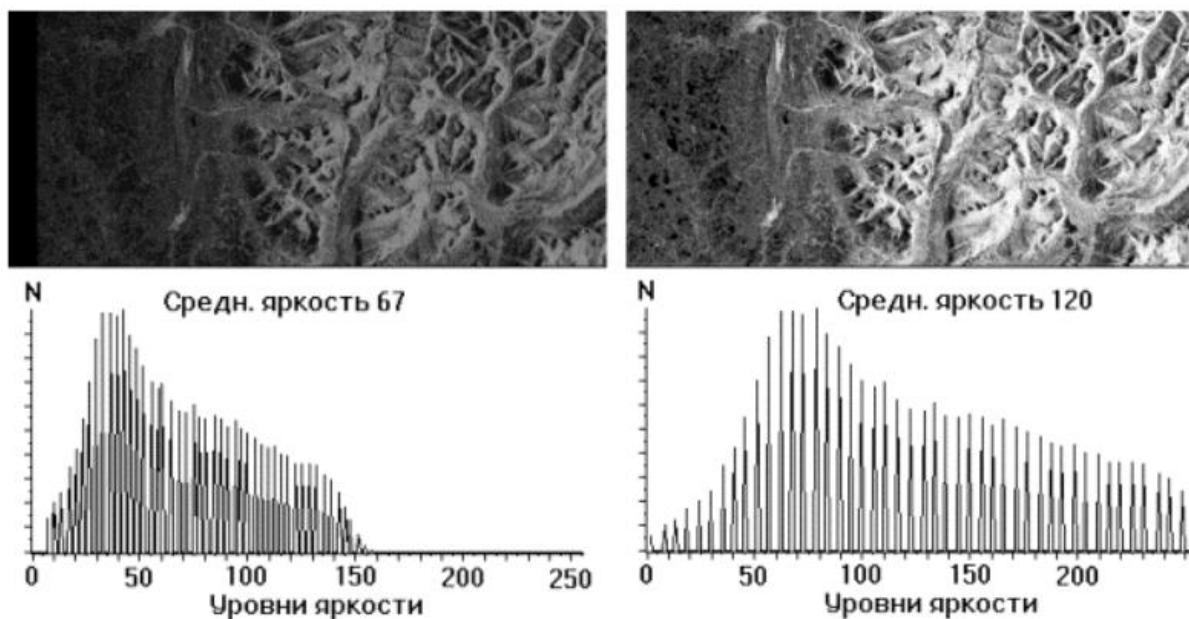


Рисунок 9 - Исходное изображение (а) и изображение после линейной растяжки гистограммы (б)

Величины для изображения, которое иллюстрирует рисунок 9а, $f_{\min}=6$, $f_{\max}=158$. Коэффициенты p и r выбираются так, чтобы $g_{\min}=0$, $g_{\max}=255$. Из формулы (1) получаем: $p=-10,01$; $r=1,67$.

Ещё больше можно усилить контрастность, применяя «нормализацию гистограммы». Данный метод подразумевает, что на весь полный диапазон уровней яркости от 0 до 255 растягивается не вся гистограмма, лежащая в промежутке от f_{\min} до f_{\max} , а её самый насыщенный фрагмент от f'_{\min} до f'_{\max} , из рассмотрения исключаются малоинформативные «хвосты». На изображении, которое иллюстрирует рисунок 10 исключено 5% точек.



а) б)
Рисунок 10 - Нормализация гистограммы

Задачей «эквализации гистограммы» (выравнивания, линеаризации) является такое преобразование, в результате которого все уровни яркости приобрели бы равную частоту, а гистограмма яркостей соответствовала бы равномерному закону распределения, что демонстрирует рисунок 11.

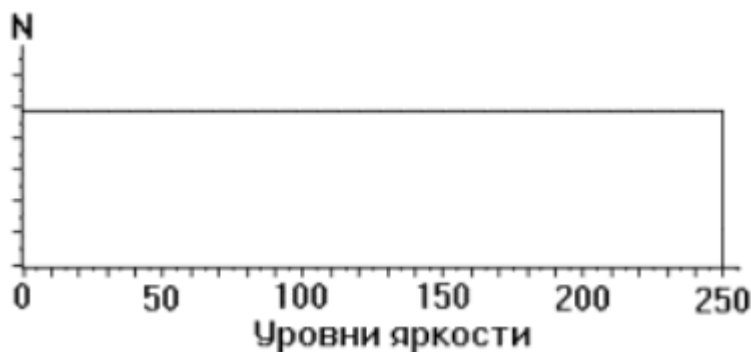


Рисунок 11 - Гистограмма, отвечающая равномерному закону распределения

Допустим, изображение содержит N точек по горизонтали и V по вертикали, а количество уровней квантования яркости равняется L . Общее число точек при этом будет равняться $N \times V$, а на каждый уровень яркости приходится, в среднем, $n_0 = NV/L$ точек. К примеру, $N=V=512$, $L=256$. В этом случае $n_0=1024$. Расстояние Δf между дискретными уровнями яркости от f_i до f_i+1 в гистограмме исходного изображения равно, но на каждый уровень приходится разное число точек. При эквализации гистограммы расстояние Δg_i между уровнями g_i и g_i+1 различно, однако число точек на каждом уровне, в среднем, одинаково и равно n_0 . Механизм эквализации достаточно прост.

Допустим, уровнями с низкой яркостью обладает небольшая доля точек, как на рис. 3а. К примеру, нулевым уровнем яркости на исходном изображении обладают 188 точек, уровнем 1 - 347 точек, уровнем 2 - 544 точки. В сумме это 1079 точек, то есть примерно n_0 . Присвоим всем этим точкам уровень 0. Пусть на исходном изображении число точек с уровнями яркости 3 и 4 в сумме приблизительно также равно n_0 . Этим точкам присваивается уровень 1. С другой стороны, пусть число точек с уровнем 45 на исходном изображении составляет 3012, то есть примерно $3n_0$. Всем этим точкам присваивается некоторый общий уровень g_i , не обязательно равный 45, а соседняя пара уровней остаются незаполненными. Описанные операции производятся со всеми уровнями яркости. Результат эквализации демонстрирует рисунок 12. Алгоритм преобразования гистограмм подбирается в каждом конкретном случае исходя из того, какой из них даст оптимальный с точки зрения решения поставленной задачи результат.

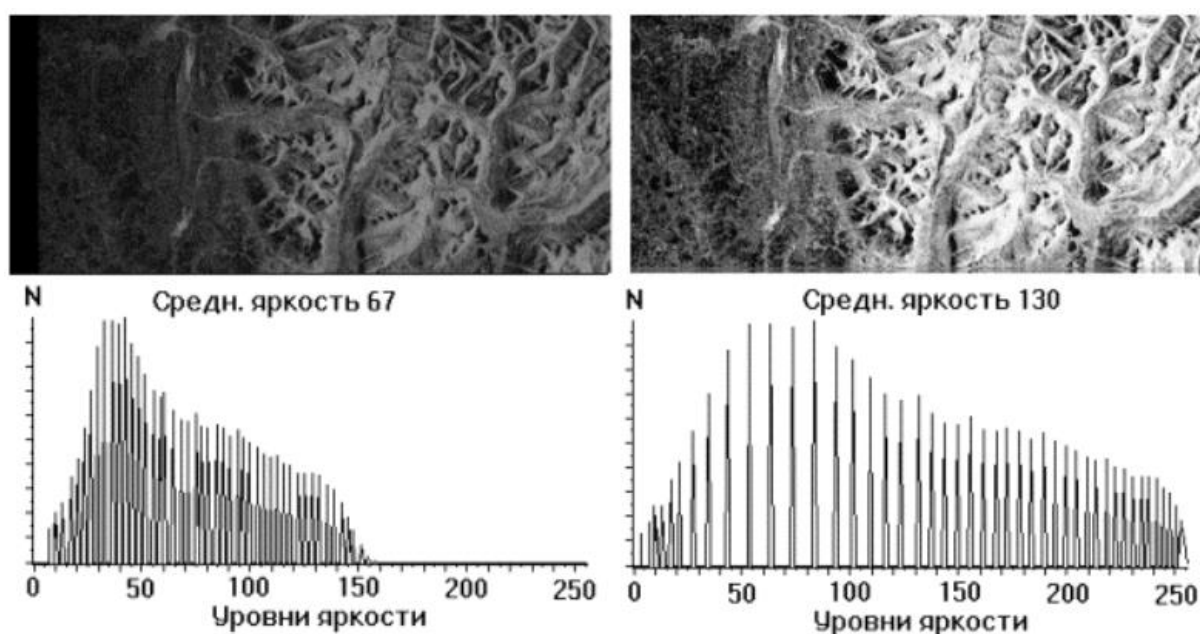


Рисунок 12 - Эквализация гистограммы

3.2 Методы фильтрации изображения

Ни одна реальная фото-, видеосистема не обеспечивает идеального качества изображений исследуемых образцов. Изображения в ходе их создания фото-, видеосистемами обычно подвергаются влиянию всевозможных случайных помех или шумов. В отличие от так называемых детерминированных искажений, зачастую описываемых поэлементными функциональными преобразованиями исходного изображения, для описания случайных помех используют модели аддитивного, импульсного и мультипликативного шумов.

Самым распространённым типом шума является статистически независимый от сигнала случайный аддитивный шум. Модель аддитивного шума употребляется тогда, когда сигнал на выходе отображающей системы или на каком-либо промежуточном шаге преобразования может рассматриваться как результирующая полезного сигнала и некоторого случайного сигнала - шума. Примерами аддитивного шума могут служить зернистость фотоплёнки, флуктуационный шум в радиотехнике, шум квантования в аналого-цифровых преобразователях и т.д.

Если влияние шума проявляется не на всей протяжённости поля изображения, а только в произвольно расположенных точках, в которых значения функции яркости заменяются некоторыми случайными величинами, то шум является импульсным. На изображении подобный шум выглядит изолированными контрастными точками. Искажённые точки равномерно распределены по всей площади изображения, а их яркость равномерно распределена в определённом диапазоне. Импульсный шум типичен для систем передачи изображений в радиоканалах с применением нелинейных методов модуляции, равно как и для цифровых систем передачи и хранения изображений. Например, импульсный шум свойственен устройствам ввода изображений с телекамеры.

Более общей линейной моделью работы с изображениями в условиях помех является модель, учитывающая помимо аддитивного шума динамические пространственные искажения. Данная модель описывает искажения, вызванные движением регистрирующей системы относительно объекта, турбулентностью атмосферы, абберрациями оптики, ошибками фокусировки и т.д.

Решению задачи фильтрации помех посвящено очень много работ, существуют различные методы и алгоритмы сглаживания шума, отличающиеся своими параметрами и выходными результатами. При сглаживании изображений используют апертуру - это прямоугольный либо квадратный участок изображения, на котором определяется функция. Заданная функция в апертуре называется весовой, или функцией окна, а сама апертура вместе с этой функцией называется маской изображения.

К настоящему времени математический аппарат теории фильтрации изображений проработан достаточно глубоко, так что задача исследователя сводится лишь к определению искажающих факторов изображения и подбору наиболее подходящего способа его улучшения. Однако, по возможности, следует стремиться к качественным снимкам уже на этапе их получения.

3.2.1 Методы линейной фильтрации

Линейные фильтры также называются сглаживающими, так как в результате их работы происходит усреднение значений точек, содержащихся в апертуре. Для этого нужно выбрать размер апертуры и специальные веса, позволяющие оказывать воздействие непосредственно на шум. В результате

работы фильтра соотношение сигнал / шум растёт пропорционально площади маски. Увеличение размера апертуры приводит к увеличению затрат вычислительной мощности для обработки изображения. Подбор коэффициентов маски улучшает подавление помех.

Фильтр Гаусса

Благодаря данному фильтру можно уменьшить с расстоянием влияние точек друг на друга. Ядро данного фильтра можно выразить формулой (2):

$$f_{gauss}(i,j) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{i^2 + j^2}{2\sigma^2}\right), \quad (2)$$

где

i, j - координаты $pxrgb$ изображения;

f_{gauss} - сигнал;

σ - шум, находящийся на исходном изображении.

Используя фильтр Гаусса, можно размывать шум, подвергая содержательные контуры изображения размытию в малой степени. К примеру, если на исходном изображении необходимо размыть мелкие детали, которые не требуют отделения от фона, а интересующие большие объекты будут выделяться на последующих этапах при помощи бинаризации.

3.2.2 Методы нелинейной фильтрации

Методы нелинейной фильтрации применяются для удаления на изображениях импульсного шума. Нелинейные фильтры используют маски и построены на основе ранговой статистики. Их алгоритм обработки изображения состоит в определении позиции каждого импульса и замене его значением оценки, в то время как значения остальных точек изображения остаются неизменными.

Медианная фильтрация

Медианная фильтрация в отличие от сглаживающей фильтрации реализует нелинейный метод подавления помех. Медианный фильтр представляет собой скользящее по площади изображения окно W , охватывающее нечётное количество отсчётов. Центральный отсчёт заменяется медианой всех элементов изображения, попавших в окно. Медианой дискретной последовательности x_1, x_2, \dots, x_L для нечётного L называют такой её элемент, для которого существуют $(L - 1)/2$ элементов, меньших или равных ему по значению, и $(L - 1)/2$ элементов, больших или равных ему по значению. Иными словами, медиана - это средний по порядку член ряда, образующегося при упорядочении исходной последовательности. К примеру, медиана ряда (18, 12, 2, 6, 4) равна 6.

Двумерный медианный фильтр с окном W определим формулой (3):

$$\hat{x}(m,n) = \underset{W}{med} y(m,n) = med[y(m+k, n+l) : (k,l) \in W]. \quad (3)$$

Как и сглаживающий фильтр, медианная фильтрация применяется для подавления аддитивных и импульсных помех на изображении. Характерной чертой медианной фильтрации, отличающей её от сглаживающей, является сохранение перепадов яркости (контуров). При этом, если перепады яркости значительны в сравнении с дисперсией аддитивного белого шума, то медианный фильтр даёт меньшее значение среднеквадратичного отклонения по сравнению с оптимальным линейным фильтром. Особенно эффективной медианная фильтрация является применительно к импульсному шуму.

Что касается импульсного шума, то, к примеру, медианная фильтрация с окном 3×3 полностью устраняет одиночные пики на равномерном фоне, а также группы из двух, трёх и четырёх импульсных пиков. В общем случае для устранения группы импульсных шумов размеры окна должны выбираться по меньшей мере вдвое больше размеров группы шумов.

Медианная фильтрация обладает высоким быстродействием и позволяет сохранить перепады оттенков, различных границ и устранить локальные пики яркости на изображениях, искажённые импульсным шумом. Медианную фильтрацию можно применять для подавления аномальных значений в массивах данных, снижения выбросов и импульсных помех благодаря свойству данных фильтров сохранять монотонные последовательности. Поэтому если сравнить медианную фильтрацию с линейной, то она лучше сохраняет контуры изображения, но, в то же время, в результате работы фильтра может возникнуть и проявиться на изображении гауссов шум. Такие фильтры также используются итеративно. Медианные фильтры достаточно хорошо работают в случаях, когда плотность шума невелика.

Ранжирующий фильтр

В основе алгоритма является ряд, подготовленный для сортировки, но он включает в себя центральный отсчёт маски фильтра. Тогда в результате сортировки этого ряда проводят вычисление значения его медианы - $med(k)$. Для того, чтобы вычислить выходное значение фильтра, используется формула (4):

$$y(k) = \alpha x(k) + (1 - \alpha)med(k), \quad (4)$$

где

α - коэффициент доверия;

$x(k)$ - значение яркости k -ой точки на исходном изображении.

3.2.3 Комбинированные и гибридные методы фильтрации

Актуальным вопросом остаётся разработка алгоритмов фильтрации, для устранения комбинированных шумов. Такая проблема решается путём

3.2.4 Адаптивные методы фильтрации

$$\bar{f}(x, y) = \frac{g(x, y) - \sigma_\eta^2((g(x, y) - m_i)}{\sigma_i^2}, \quad (5)$$

$f(x, y)$ - отфильтрованное изображение;
 $g(x, y)$ - зашумлённое изображение.

						ДП-230102.65-001019120ПЗ	Лист
							35
Изм.	Колич.	Лист.	№ док	Подпись	Дата		

3.3 Методы бинаризации изображения

Для последующего изучения объектов изображения их выделяют относительно фона или посторонних элементов. Зачастую выбор способа выделения объектов зависит как от способа получения изображений, так и от типа объекта исследования. Сегментация, то есть разбиение изображения на участки, для которых справедлива определённая мера однородности, например, выделение на изображении участков примерно равной яркости, является одной из основных и одновременно одной из самых сложных среди комплекса задач обработки изображений. Центральным методом сегментации является пороговая обработка. Это один из наиболее важных способов поэлементного преобразования, который почти всегда применяется перед непосредственным процессом распознавания изображений. Эта операция заключается в сравнении яркости каждой точки изображения с установленным порогом. Выбор соответствующего значения порога предоставляет возможность выделения на исходном изображении зон определённого вида.

Процедура порогового разделения, в результате которой получается бинарное изображение, получила название «бинаризация». Целью процедуры бинаризации является кардинальное сокращение количества информации, с которой приходится работать. Главным параметром такого преобразования является порог t , с которым сравнивается яркость каждой точки многоградиационного изображения. По результатам сравнения, точке присваивается значение 1 или 0. Так как количество информации в бинарном изображении почти на порядок меньше, чем в исходном многоградиационном, то бинарное изображение заметно проще хранить и обрабатывать. Качественная бинаризация значительно упрощает последующую обработку изображения. Однако, изъяны в алгоритме бинаризации могут привести к искажениям (разрывы в границах, утрата значащих деталей, нарушение целостности объектов, шум и т.д.). Различные методы бинаризации имеют свои слабые места: поэтому каждый метод должен быть применён в своей области.

Бинаризация с верхним (6) или нижним (7) считается наиболее простой, поскольку в ней применяется единственное значение порога:

$$f'(m,n) = \begin{cases} 0, f(m,n) \leq t \\ 1, f(m,n) > t \end{cases} \quad (6)$$

$$f'(m,n) = \begin{cases} 0, f(m,n) \geq t \\ 1, f(m,n) < t \end{cases} \quad (7)$$

Все значения меньше или равные t преобразуются в 1, больше t - в 0 при бинаризации с верхним порогом, и наоборот - с нижним.

Для выделения участков, в которых значения яркости точек может меняться в определённом диапазоне, применяется бинаризация с двойным ($t_1 < t_2$) порогом (8):

$$f'(m, n) = \begin{cases} 0, f(m, n) \geq t_1 \\ 1, t_1 < f(m, n) \leq t_2 \\ 0, f(m, n) > t_2 \end{cases} \quad (8)$$

Неполная пороговая обработка (9) позволяет получить изображение, которое значительно удобнее для последующей обработки, так как лишено фона, но сохраняет все детали, имеющиеся на исходном изображении.

$$f'(m, n) = \begin{cases} f(m, n), & f(m, n) > t \\ 1, & f(m, n) \leq t \end{cases} \quad (9)$$

Результатом многоуровневого порогового преобразования (10) является изображение, не являющееся бинарным, но состоящее из участков с различной яркостью.

$$f'(m, n) = \begin{cases} 1, f(m, n) \in D_1 \\ 2, f(m, n) \in D_2 \\ \dots \\ n, f(m, n) \in D_n \\ 0, \text{ иначе} \end{cases} \quad (10)$$

Бинаризация может быть глобальной, применяемой для всего изображения и локальной, захватывающей лишь часть изображения.

3.3.1 Методы локальной бинаризации

Метод Оцу

Самым эффективным по качеству и по быстродействию принято считать метод Оцу. При помощи этого метода вычисляется порог t , минимизирующий среднюю ошибку сегментации, то есть среднюю ошибку от принятия решения о принадлежности точек изображения объекту или фону. Метод использует гистограмму распределения значений яркости точек растрового изображения. Значения яркостей точек изображения можно рассматривать как случайные величины, а их гистограмму - как оценку плотности распределения вероятностей. Гистограмма формируется по значениям $p_i = k_i / K$, где K - это общее количество точек на изображении, k_i - это количество точек с уровнем яркости i . Диапазон яркостей делится на две группы с помощью порогового значения уровня яркости t (целое значение от 0 до K). Каждой группе соответствуют относительные частоты ω_0 (11) и ω_1 (12):

$$\omega_1(t) = \sum_{i=1}^t p_i \quad (11)$$

$$\omega_2(t) = \sum_{i=t+1}^K p_i = 1 - \omega_1(t) \quad (12)$$

Средние для каждой группы изображения уровни μ_1 (13) и μ_2 (14):

$$\mu_1(t) = \sum_{i=1}^t \frac{ip_i}{\omega_1} \quad (13)$$

$$\mu_2(t) = \sum_{i=t+1}^K \frac{ip_i}{\omega_2} \quad (14)$$

Затем определяется максимальное значение оценки качества разделения изображения на две группы (15):

$$\eta(t) = \max_{t-1}^{K-1} \left(\frac{\sigma_{\text{мг}}^2(t)}{\sigma_{\text{од}}^2} \right), \quad (15)$$

где

$\sigma_{\text{мг}}^2 = \omega_1 \omega_2 (\mu_2 - \mu_1)^2$ - межгрупповая дисперсия;

$\sigma_{\text{од}}^2$ - общая дисперсия изображения в целом.

Определение порога на основе градиента яркости изображения

Метод основывается на предположении, что исходное изображение можно разбить на два класса - фон и объекты. Процедура расчёта порогового значения состоит из двух этапов. Сначала вычисляется модуль градиента яркости для каждой точки изображения (16):

$$P(i, j) = \max\{|P_i(i, j)|, |P_j(i, j)|\}, \quad (16)$$

где

$$P_i(i, j) = f(i + 1, j) - f(i - 1, j);$$

$$P_j(i, j) = f(i, j + 1) - f(i, j - 1).$$

На втором этапе алгоритма производится определение порога t по формуле (17):

$$t = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} f(i, j) P(i, j)}{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} P(i, j)} \quad (17)$$

3.3.2 Методы глобальной бинаризации

Метод Бернсена

Метод Бернсена нередко применяется для схематических и картографических изображений. Входное изображение делится на квадраты $g \times g$ (g - нечётное) с центром в точке (m, n) . Апертура пробегает в цикле по всем точкам входного изображения. На каждом этапе определяются минимум и максимум. Далее вычисляется среднее значение как половина суммы максимума и минимума. Если яркость текущей точки превышает среднее и меньше ε (константа, заданная пользователем) - она становится белой, иначе - чёрной. Если среднее значение меньше порога контраста - то текущая точка становится того цвета, который задавался параметром «цвет сомнительной точки». Исследования показывают, что наилучшие результаты достигаются для значений $\varepsilon=15$ и $g=15$.

У рассматриваемого метода присутствует ряд минусов: при обработке монотонных областей яркости могут формироваться значительные паразитные помехи, в некоторых случаях он приводит к появлению ложных черных областей. Недостатки могут быть компенсированы с помощью дополнительной обработки - постпроцессинга. Метод считается более быстрым, чем метод Оцу, даже в совокупности с этапом постпроцессинга.

Метод Эйквила

Одним из самых производительных методов бинаризации принято считать метод Эйквила. Он часто используется при работе с контрастными и чёткими изображениями. Входное изображение обрабатывается двумя концентрическими окнами: маленьким L , и большим V . Чаще всего выбирается квадратная форма окон. Последовательно слева направо и сверху вниз с шагом, равным стороне маленького окна L , окна накладываются на изображение. Для окна V рассчитывается порог t таким образом, чтобы разделить точки на две группы. Если матожидания уровня яркости в двух группах имеют разницу, превышающую некоторый заданный пользователем уровень $|\mu_1 - \mu_2| \geq 1$, то все точки внутри окна L бинаризируются в соответствии с порогом t . Иначе, яркость точек из окна L заменяется некоторым близким значением.

Метод Ниблэка

Смысл данного метода заключается в изменении порога бинаризации t от точки к точке на основании локальных значений среднего и среднеквадратичного отклонений. Порог яркости в точке (n, m) рассчитывается следующим образом (18):

$$t(n, m) = \alpha(n, m) + l \cdot \beta(n, m), \quad (18)$$

где

$\alpha(n, m)$ - среднее отклонение выборки для окрестности точки;

β (n, m) - среднеквадратичное отклонение выборки для окрестности точки.

Размер окрестности должен быть минимальным, но таким, чтобы сохранить локальные детали изображения. Вместе с тем он должен быть достаточно большим, чтобы снизить влияние помех на результат. Значение β указывает, какая часть границы объекта принимается в качестве самого объекта. Значение β равное минус 0,2 задаёт достаточно неплохое разделение объектов, если они представлены черным цветом, а значение β равное 0,2, - если объекты представлены белым цветом.

Метод Яновица и Брукштейна

В качестве пороговой поверхности бинаризации используется поверхность потенциалов, формируемая на основе локальной максимизации градиента яркости. Значение градиента яркости часто рассчитывается при помощи контурных операторов Собеля или Кэнни. Изображение фильтруется для получения контурных линий толщиной в 1 точку, а после усредняющим фильтром 3×3 . Потенциальная поверхность после этого строится по итерационной интерполирующей схеме. Вычисление поверхности проводится, начиная от контурных точек. Для каждой точки, не являющейся контурной, вычисляется интерполяционный остаток $R(x, y)$ и новое значение точки $P(x, y)$ на $n+1$ -ом шаге должно вычисляться в соответствии с формулами (19) и (20):

$$P_{n+1}(x, y) = \frac{P_n(x, y) + \beta \cdot R_n(x, y)}{4}, \quad (19)$$

$$R_n(x, y) = P_n(x - 1, y) + P_n(x + 1, y) + P_n(x, y - 1) + P_n(x, y + 1) - 4 \cdot P_n(x, y), \quad (20)$$

где β в пределах $1 \leq \beta \leq 2$ для быстрой сходимости.

4 Реализация системы параметрического анализа клеточных структур деревьев хвойных пород

4.1 Описание входных данных системы

В качестве входных данных системы выступают фотографические изображения клеточных структур деревьев хвойных пород, полученные при помощи электронного микроскопа. Рисунок 13 иллюстрирует примеры таких фотографий. Размер изображений - 1388×1040 точек, формат 8 bit JPEG (оттенки серого).

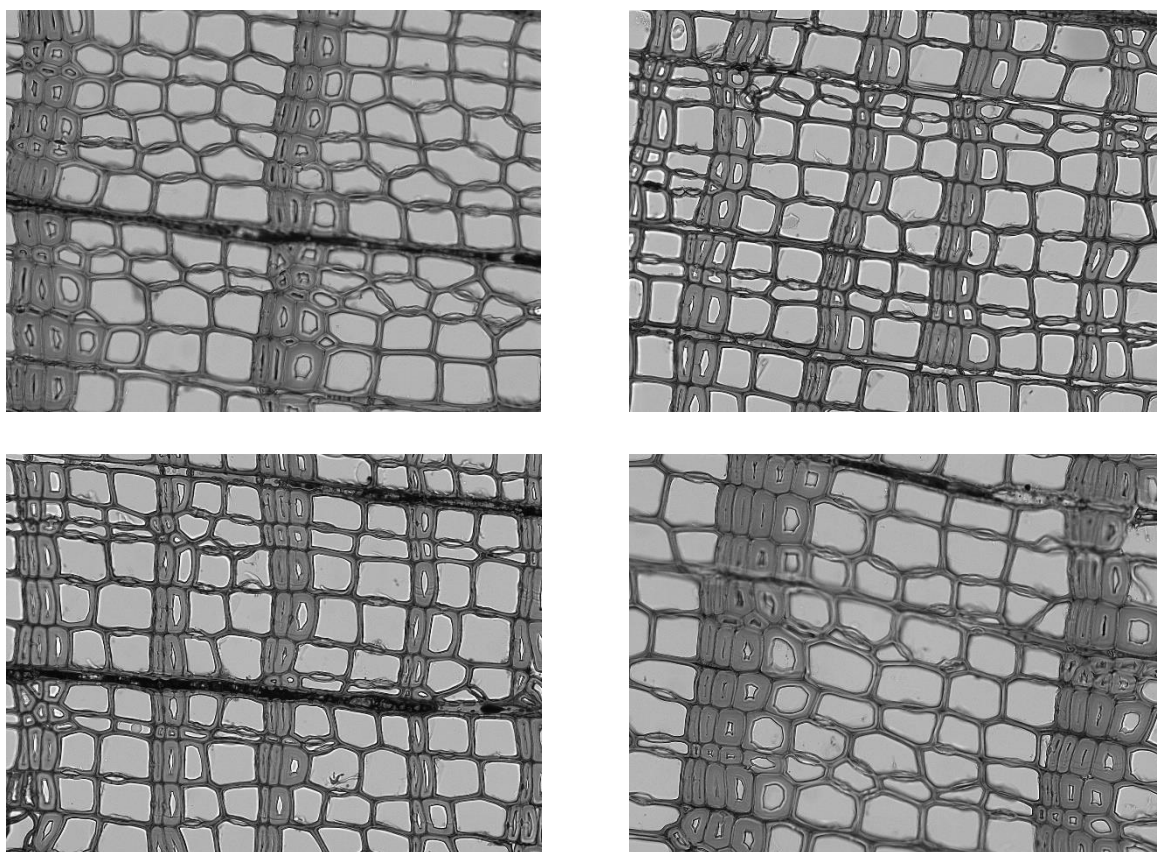


Рисунок 13 - Примеры изображений для обработки

Особенностью входных данных является большое количество дефектов, которые значительно усложняют, а зачастую делают невозможной их обработку. Дефекты вызваны плохим качеством исследуемых образцов - разрывы стенок, посторонние предметы, загибание краёв образца, бликующие капли на поверхности и т.п., что иллюстрирует рисунок 14. Также другая часть дефектов связана с ошибками в процессе микросъёмки - смещения по линиям склейки, пустые области в кадре, отсутствие резкости на участках исследуемого образца, замутнённые регионы, что наглядно демонстрирует рисунок 15. В ходе работы была выявлена необходимость предварительной сортировки и, возможно, дополнительной предобработки в подобных случаях,

Изм.	Колич.	Лист.	№ док	Подпись	Дата

ДП-230102.65-001019120ПЗ

Лист

41

иначе смысл автоматического распознавания теряется. В рамках описываемой разработки подобный функционал не реализовывался.

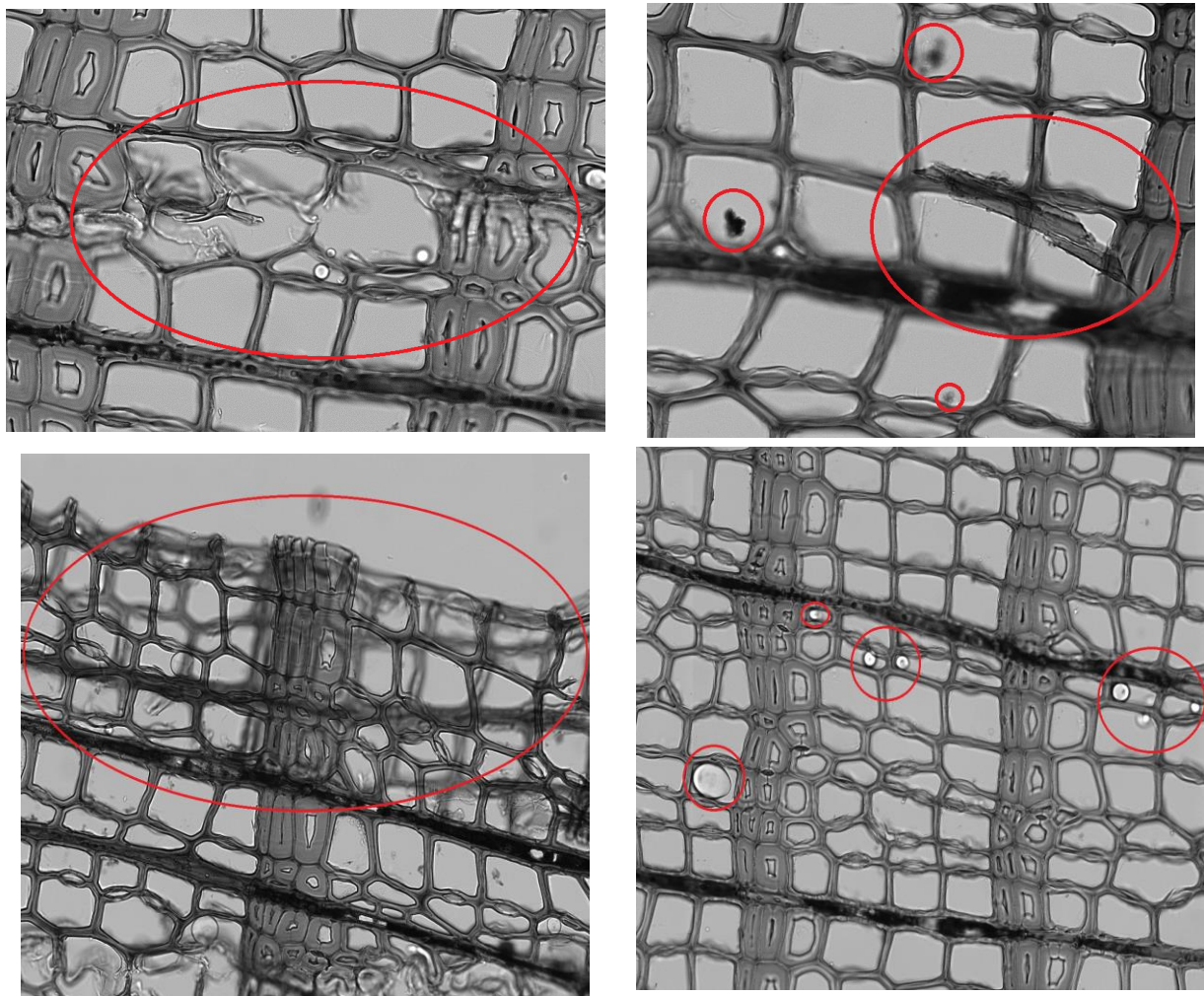


Рисунок 14 - Примеры дефектов исследуемых образцов

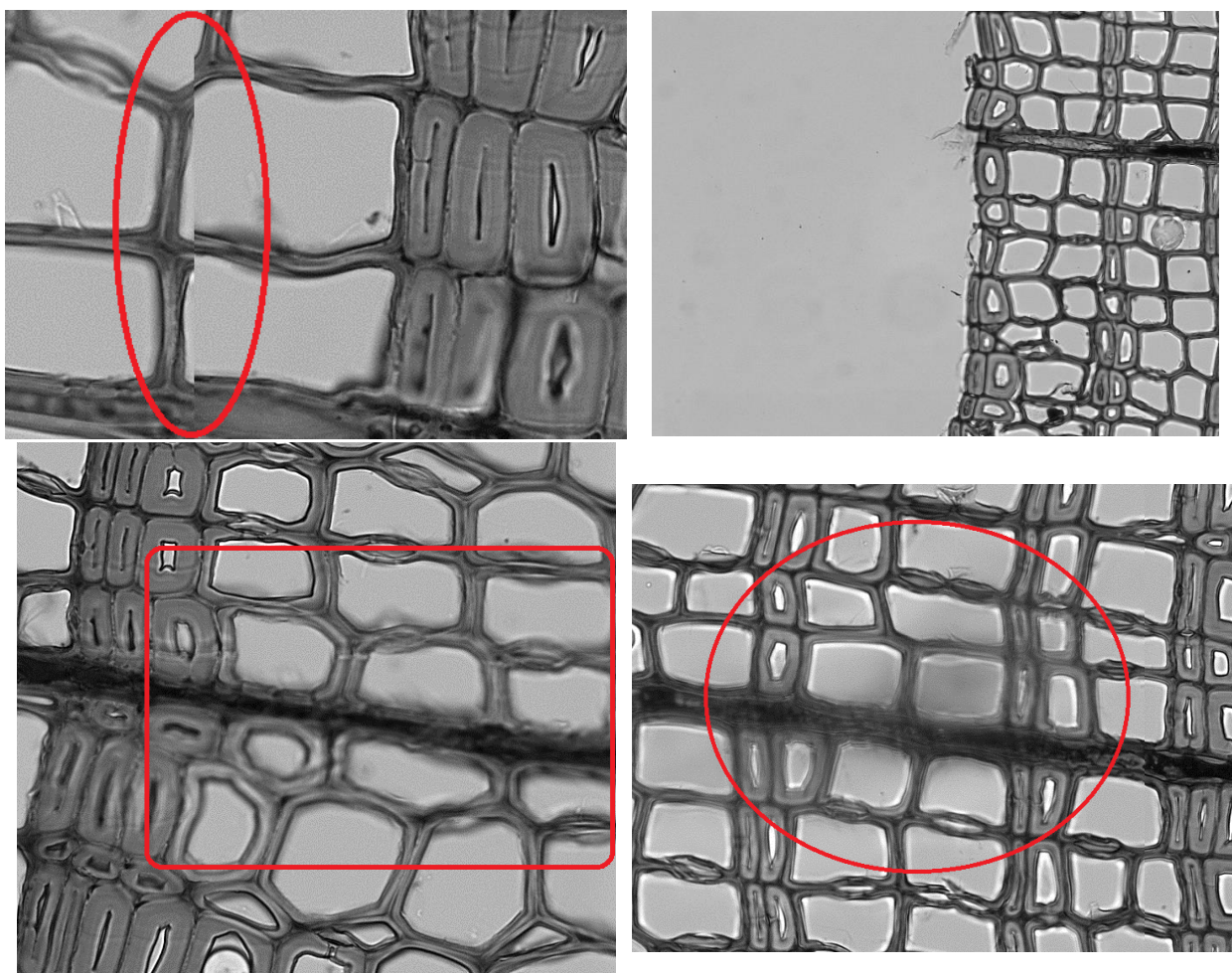


Рисунок 15 - Примеры дефектов фотоизображения

4.2 Состав модулей системы

Разработанный прототип системы автоматизированного параметрического анализа клеточных структур деревьев хвойных пород состоит из трёх модулей: Main.pas, class_TAnalyzePicThread.pas и LoadPic.pas.

Модуль Main и привязанная к нему форма Main.dfm содержит графический интерфейс пользователя и процедуры работы с ним.

В модуле class_TAnalyzePicThread реализован весь основной функционал проекта. При разработке было принято решение вынести всю работу по анализу входных изображений в отдельный от интерфейса пользователя поток, для чего был создан класс TAnalyzePicThread, являющийся потомком класса TThread. При загрузке объекта анализа создаётся экземпляр класса, выполняющий все необходимые вычислительные процедуры, при этом пользовательский интерфейс остаётся доступен для управляющих воздействий. Более подробно функционал модуля описан в нижеследующих разделах.

Модуль LoadPic включает в себя процедуры загрузки входных изображений и приведение их формата к формату класса TBitmap. Выделение в отдельный модуль обусловлено тем, что, во-первых, его процедуры не связаны напрямую с основным функционалом проекта, а во-вторых, для возможности независимого расширения списка поддерживаемых форматов при необходимости. Работа модуля описана в п. 4.3.1.

4.3 Предварительная обработка изображения

4.3.1 Загрузка изображений

Процедура загрузки изображения унифицирована и позволяет работать не только с файлами формата JPEG, но также с форматами BMP, GIF и PNG. Тип входного изображения определяется автоматически, и в соответствии с ним производится преобразование в формат TBitmap. Данный функционал реализован в процедурах LoadPicToBitmap, GetPicFormat, GIFtoBMP, PNGtoBMP и JPEGtoBMP модуля LoadBitmap.

Процедура LoadPicToBitmap осуществляет загрузку входного файла в поток TMemoryStream, вызов алгоритма определения формата изображения и выбор метода его преобразования к TBitmap:

```
procedure LoadPicToBitmap(var btmp: TBitmap; const fileName:
TFileName);
var
  picStream: TMemoryStream;
begin
  picStream := TMemoryStream.Create;
  try
    picStream.LoadFromFile(fileName);
    case GetPicFormat(picStream.Memory) of
      BMP:
        btmp.LoadFromStream(picStream);
      JPG:
        JPEGtoBMP(btmp, picStream);
      PNG:
        PNGtoBMP(btmp, picStream);
      GIF:
        GIFtoBMP(btmp, picStream);
    end;
    btmp.PixelFormat := pf8bit;
  finally
    picStream.Free;
  end;
end;
```

Функция GetPicFormat производит определение формата входного изображения по данным заголовка:

```
function GetPicFormat(start: Pointer): TImageType;
begin
```



```
begin
  p := ap_bmp.ScanLine[y];
  for x := 0 to cntrX do
    ap_picArr[x, y] := p^x;
  end;
end;
```

4.3.2 Повышение контрастности изображений

Исходные изображения, поступающие на вход системы, отличаются низкой контрастностью, которая существенно затрудняет процесс последующего распознавания объектов. Уровни яркости пикселей сосредоточены в диапазоне от 64 до 192, при этом чисто чёрные и чисто белые цвета отсутствуют практически полностью. Опытным путём было установлено, что наилучшие результаты даёт двухпроходное повышение контрастности с промежуточной медианной фильтрацией. Рисунок 16 демонстрирует результат первого прохода алгоритма, рисунок 17 - второго (после медианной фильтрации).

Алгоритм повышения контрастности состоит из двух этапов. На первом составляется массив яркостей, рассчитанный в зависимости от переданного в процедуру коэффициента, который, в свою очередь, либо задаётся пользователем, либо определяется автоматически методом Оцу (описание реализации метода приведено в п. 4.3.4). На втором происходит замена яркостей пикселей изображения на значения подготовленного массива.

Исходный код подготовки массива яркостей, реализующий правило преобразования исходных значений в модифицированные:

```
var
    x, k: Integer;
    mult: single;
begin
    mult := 1 + ap_ContrastValue / 100;
    for x := 0 to 255 do
        begin
            k := Round((x - treshhold) * mult) + treshhold;
            if k > 255 then
                c_ColorArray[x] := 255
            else if k < 0 then
                c_ColorArray[x] := 0
            else
                c_ColorArray[x] := k;
            end;
        end;
    end;
```

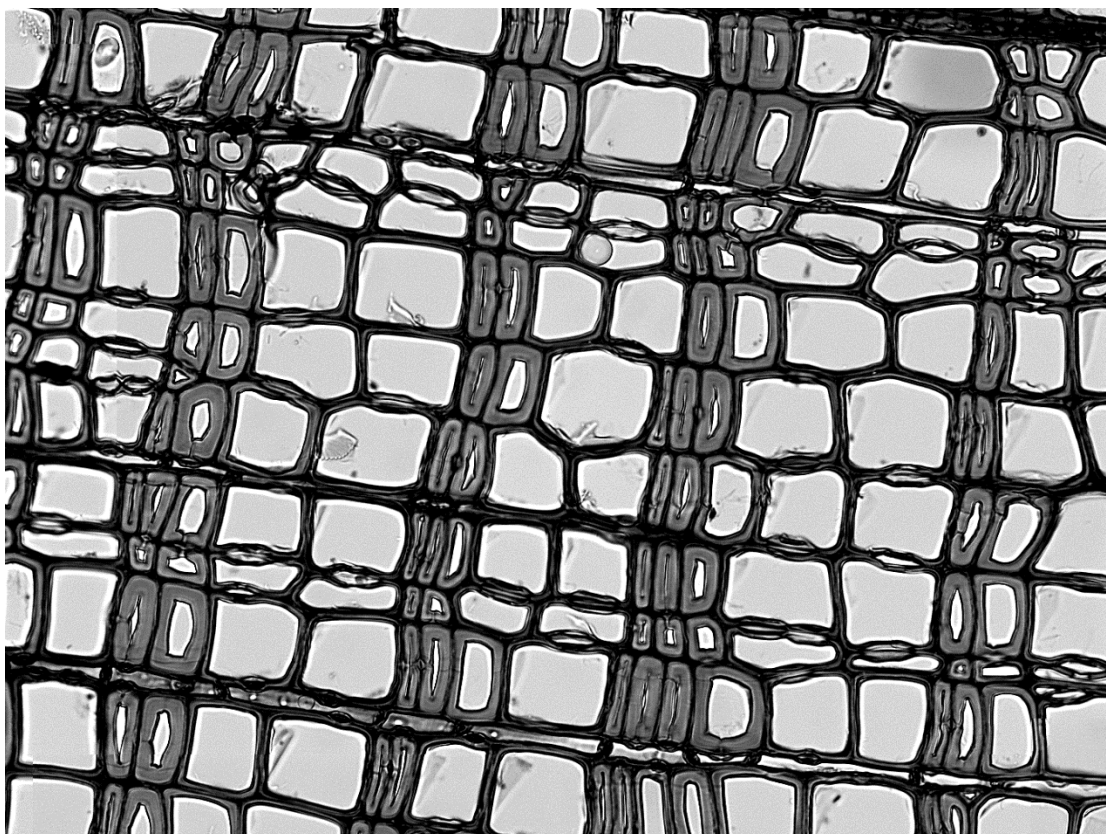


Рисунок 16 - Первый проход алгоритма повышения контрастности

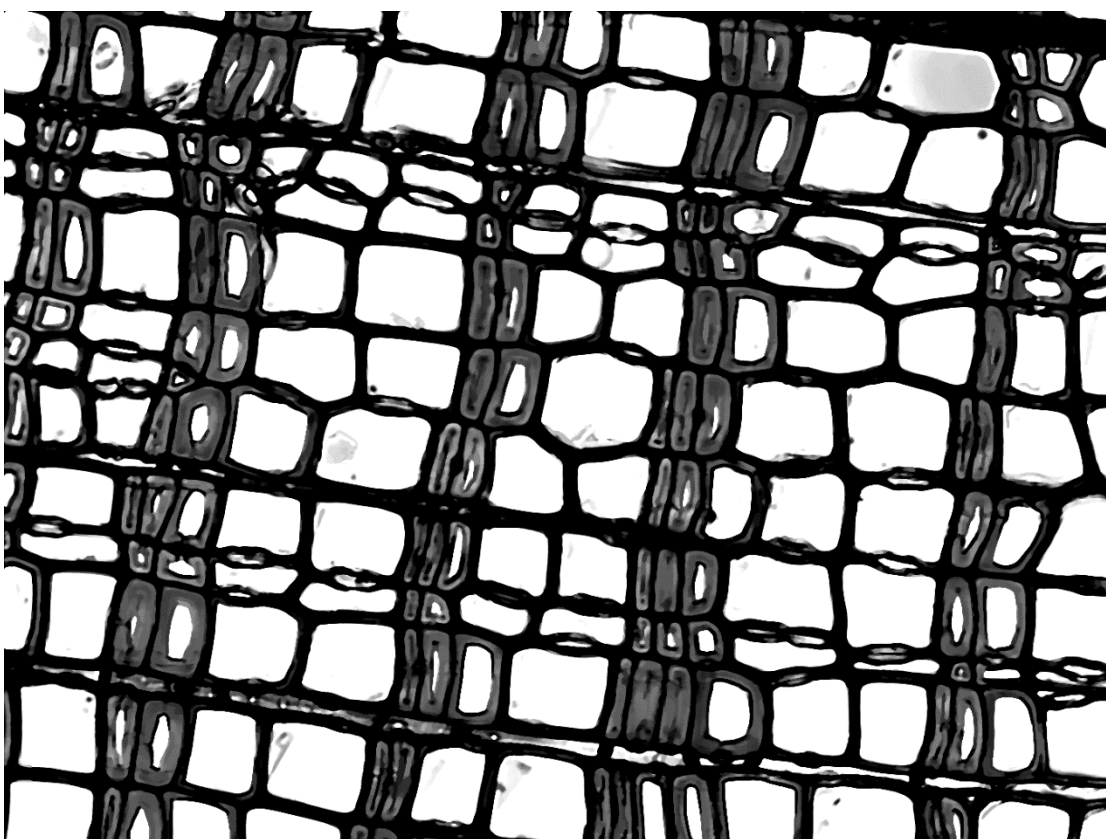


Рисунок 17 - Второй проход алгоритма повышения контрастности

Изм.	Колич.	Лист.	№ док	Подпись	Дата

ДП-230102.65-001019120ПЗ

Лист

47

Исходный код применения вычисленных значений яркостей к изображению:

```
var
  x, y, cntrX, cntrY: Integer;
begin
  GetColorArray(GetOtsuThreshold);
  cntrX := ap_PicWidth - 1;
  cntrY := ap_PicHeight - 1;
  for y := 0 to cntrY do
    for x := 0 to cntrX do
      ap_picArr[x, y] := c_ColorArray[ap_picArr[x, y]];
    end;
  end;
```

Результатом работы алгоритма повышения контрастности является изображение с максимально отделёнными друг от друга по диапазону яркостей люменами и стенками клеток.

4.3.3 Медианная фильтрация

В целях снижения импульсного шума и выравнивания локальных перепадов яркостей для предобработки применена медианная фильтрация между двумя проходами повышения контрастности, результат которой иллюстрирует рисунок 18. В качестве параметров фильтра в процедура принимает размер апертуры и вес центрального элемента. Апертура определяет размер скользящего окна, осуществляющего сглаживание, а вес центрального элемента - то, сколько раз текущая точка включается в массив сортировки. Сортировка может быть осуществлена любым известным способом, а в текущей реализации применён алгоритм быстрой сортировки.

Изм.	Колич.	Лист.	№ док	Подпись	Дата

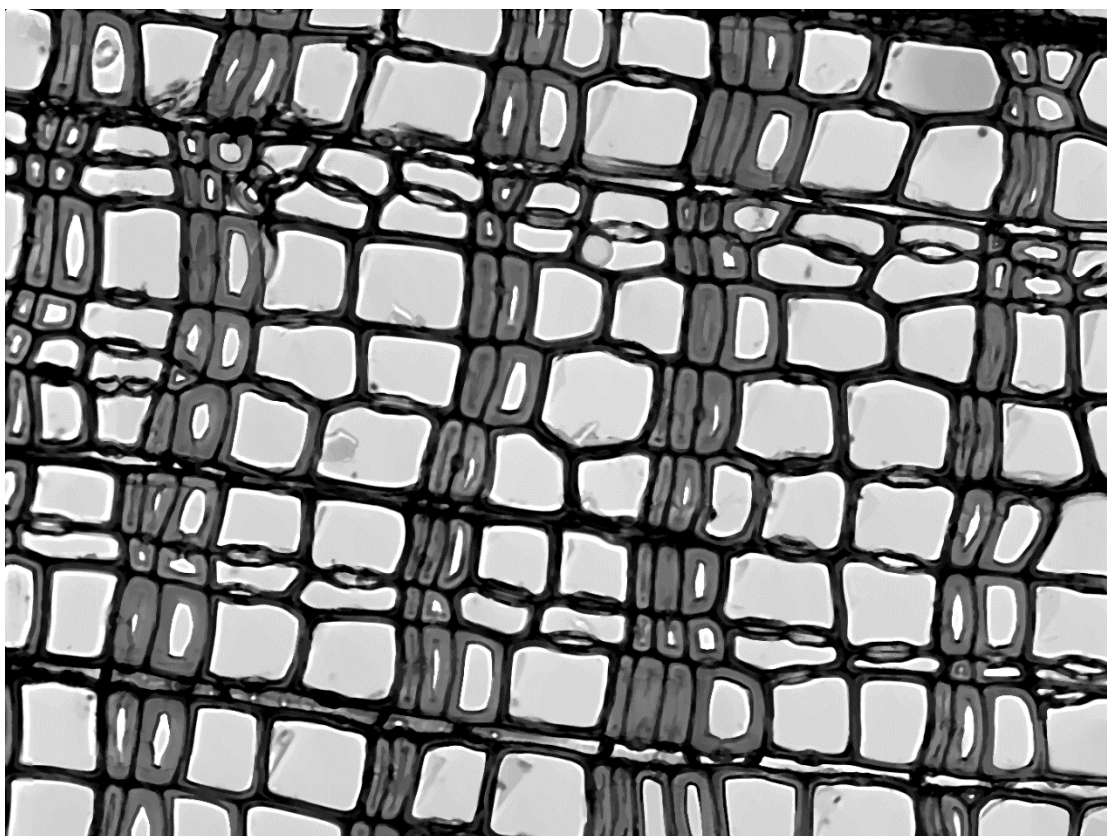


Рисунок 18 - Результат медианной фильтрации

Исходный код процедуры медианной фильтрации:

```

procedure TAnalyzePicThread.DoMedianFilter;
var
  arrLength: Byte;
  pixelArray: array of Byte;
  i, cntrV, v, cntrH, h, cntr, vB, hB, vBMax, vBMin, hBMax, hBMin:
  Integer;
begin
  arrLength := (ap_MFAperture * 2 + 1) * (ap_MFAperture * 2 + 1) +
    ap_MFCenterNum;
  SetLength(pixelArray, arrLength);

  cntrV := ap_PicHeight - 1;
  cntrH := ap_PicWidth - 1;
  for v := 0 to cntrV do
    begin
      for h := 0 to cntrH do
        begin
          // Заносим все пиксели окошка в массив
          cntr := 0;
          vBMin := v - ap_MFAperture;
          vBMax := v + ap_MFAperture;
          for vB := vBMin to vBMax do
            begin
              hBMin := h - ap_MFAperture;
              hBMax := h + ap_MFAperture;
              for hB := hBMin to hBMax do
                begin

```

Изм.	Колич.	Лист.	№ док	Подпись	Дата

ДП-230102.65-001019120ПЗ

Лист

49

```

        if ((hB >= 0) and (hB <= cntrH)) and ((vB >= 0) and (vB <=
cntrV))
        then
            pixelArray[cntr] := ap_picArr[hB, vB]
        else
            pixelArray[cntr] := 0;
            inc(cntr);
        end;
    end;

    // Добавляем текущий элемент в массив (вес центра)
    for i := 0 to ap_MFCenterNum - 1 do
        pixelArray[High(pixelArray) - i] := ap_picArr[h, v];

    // Сортируем массив
    SortArr(pixelArray, 0, arrLength - 1);
    // Берём то, что посередине и присваиваем текущему пикселю
    ap_picArr[h, v] := pixelArray[arrLength div 2];
    end;
end;
end;

```

Экспериментально было установлено, что наилучшие результаты для последующего распознавания имеющихся входных изображений даёт медианная фильтрация с апертурой 5 и весовым коэффициентом 1. Именно эти значения установлены по умолчанию в интерфейсе пользователя.

4.3.4 Бинаризация

Для выделения на предварительно обработанном изображении люменов клеток в ходе разработки системы было определено, что наилучшие результаты даёт адаптивная бинаризация. В ходе неё массив точек подвергается операции порогового разделения - все точки с яркостью ниже порога преобразуются в чёрные, выше порога - в белые. Рисунок 19 демонстрирует результат бинаризации для выделения люменов. Исходный код процедуры бинаризации:

```

procedure TAnalyzePicThread.dsBinarizeBMP(threshold: Byte);
var
    x, cntrX, y, cntrY: Integer;
begin
    cntrX := ap_PicWidth - 1;
    cntrY := ap_PicHeight - 1;
    for y := 0 to cntrY do
        begin
            for x := 0 to cntrX do
                // Если больше, чем порог
                if (ap_picArr[x, y] > threshold) then
                    ap_picArr[x, y] := 255 // То цвет белый
                else
                    ap_picArr[x, y] := 0; // Иначе цвет черный
                end;
            end;
        end;
    end;
end;

```

end;

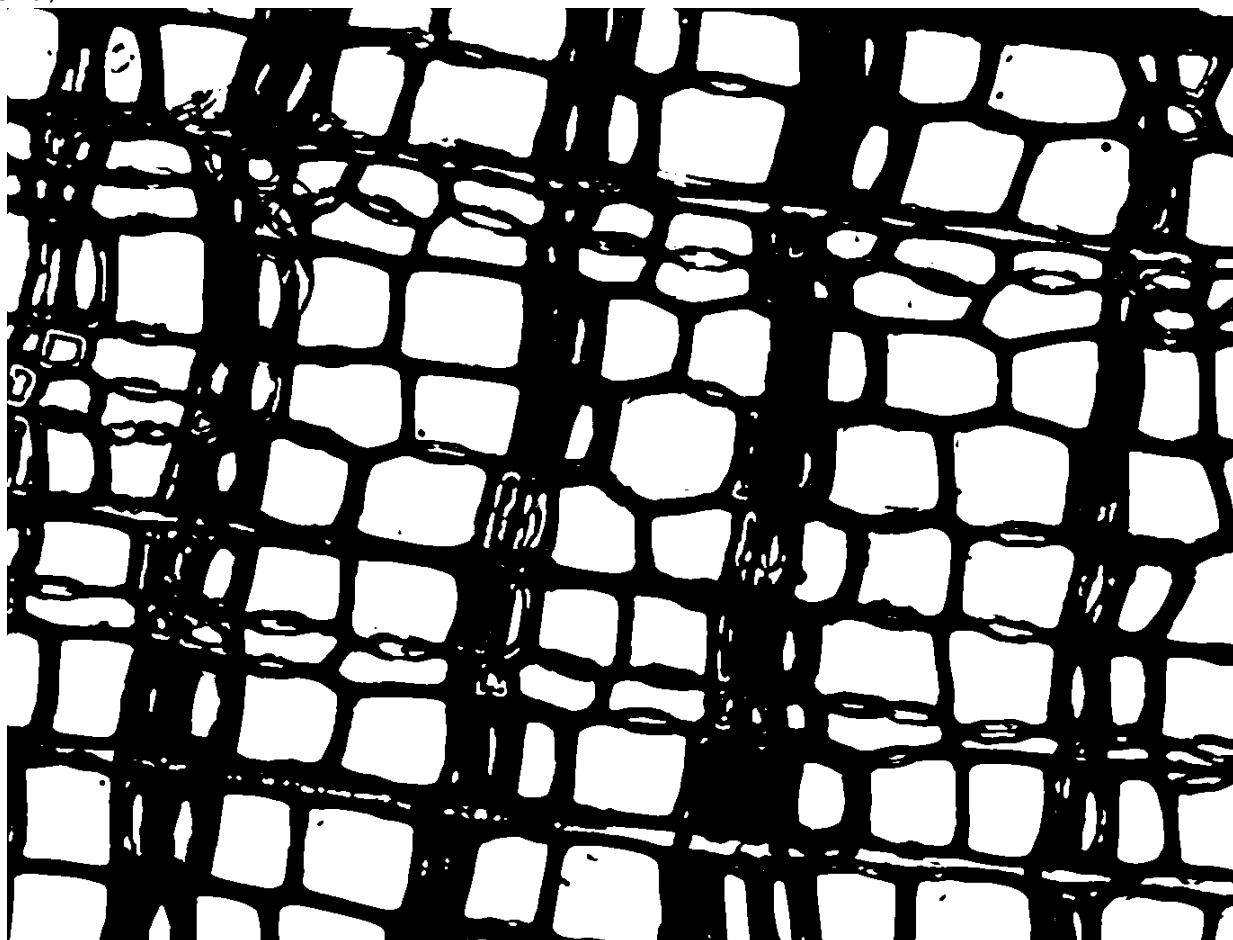


Рисунок 19 - Результат бинаризации изображения для выделения люменов

Для определения порога бинаризации применён метод Оцу, позволяющий реализовать адаптивный подход для каждого входного изображения. Реализованный алгоритм определения порога методом Оцу:

```
function TAnalyzePicThread.GetOtsuThreshold: Byte;
var
  min, max: Byte;
  x, y, cntrX, cntrY, cntr, i, k: Integer;
  pic1Arr: array of Byte;
  hist: array of Integer;
  sigma, w1, a, alpha1, beta1, m, n, maxSigma: Double;
begin
  // Перевод двумерного массива изображения в одномерный
  SetLength(pic1Arr, ap_PicWidth * ap_PicHeight);
  k := 0;
  cntrX := ap_PicWidth - 1;
  cntrY := ap_PicHeight - 1;
  for x := 0 to cntrX do
    begin
      for y := 0 to cntrY do
        begin
          pic1Arr[k] := ap_picArr[x, y];
          inc(k);
        end;
      end;
    end;
```

Изм.	Колич.	Лист.	№ док	Подпись	Дата

ДП-230102.65-001019120ПЗ

Лист

51

```

end;

// Посчитаем max и min яркость точек
min := piclArr[0];
max := piclArr[0];
cntr := High(piclArr);
for i := 1 to cntr do
begin
    if (piclArr[i] < min) then
        min := piclArr[i]
    else if (piclArr[i] > max) then
        max := piclArr[i];
end;

// Создание гистограммы, ограниченной значениями min и max
SetLength(hist, max - min + 1);

// Заполнение гистограммы 0
cntr := high(hist);
for i := 0 to cntr do
    hist[i] := 0;

// Вычисление высоты бинов
cntr := High(piclArr);
for i := 0 to cntr do
    inc(hist[piclArr[i] - min]);

// Введение двух вспомогательных чисел:
m := 0; // m - сумма высот бинов * на положение их середины
n := 0; // n - сумма высот бинов
cntr := max - min;
for i := 0 to cntr do
begin
    m := m + i * hist[i];
    n := n + hist[i];
end;

maxSigma := -1; // Max значение межклассовой дисперсии
result := 0; // Порог, соответствующий maxSigma

alpha1 := 0; // Сумма высот бинов для 1 класса
// Сумма высот бинов для 1 класса * на положение их середины
beta1 := 0;
// alpha2 не нужна, так как она = m - alpha1
// beta2 не нужна, так как она = n - alpha1

// Проход по всем возможным значениям порога
for i := 0 to cntr - 1 do
begin
    alpha1 := alpha1 + i * hist[i];
    beta1 := beta1 + hist[i];

    // Расчёт вероятности 1 класса
    w1 := beta1 / n;
    // w2 не нужна, она = 1 - w1

    // A = a1 - a2, где a1, a2 - средние арифметические классов 1 и 2

```

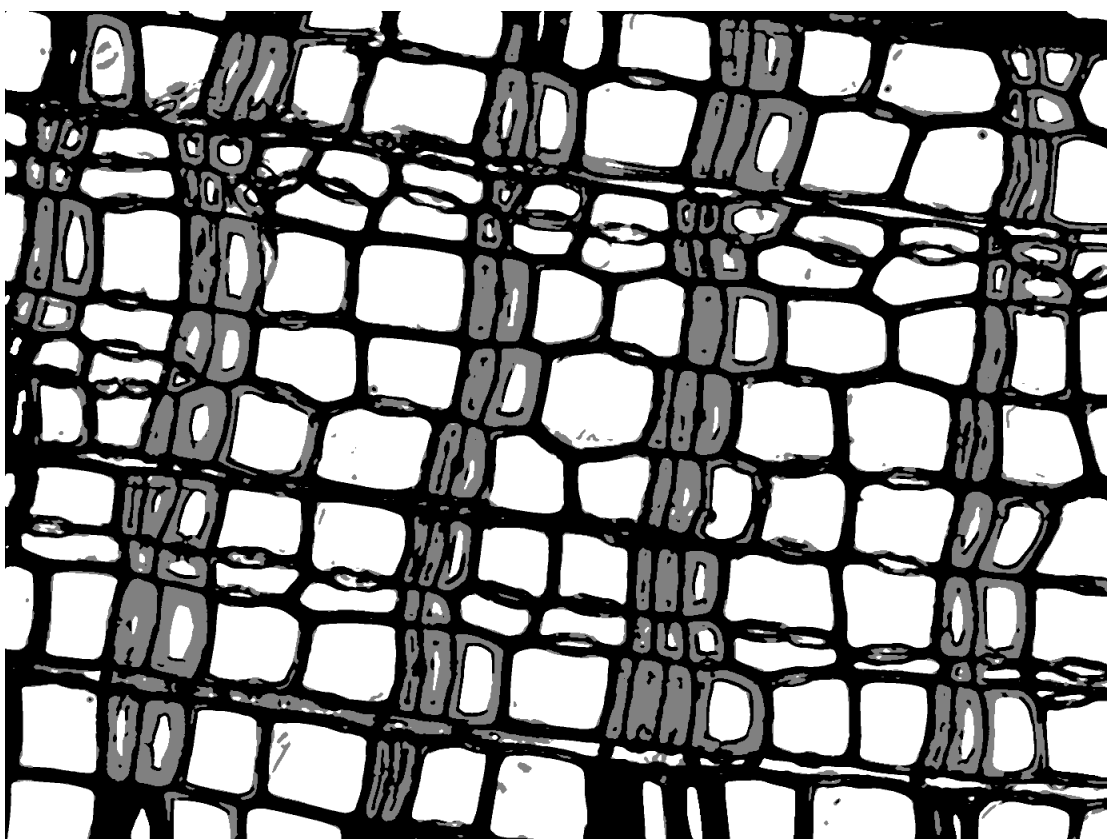
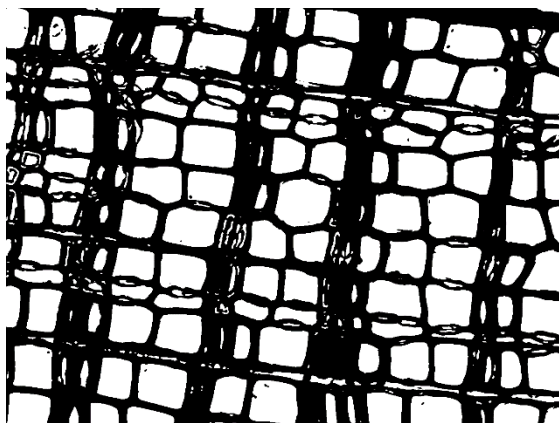



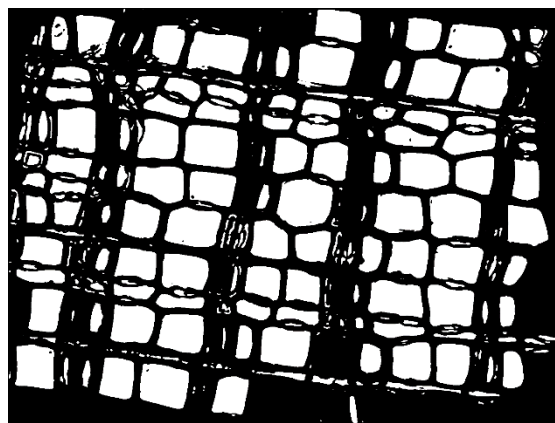
Рисунок 20 - Бинаризация с двойным порогом массива «толстых» стенок

4.3.5 Удаление нецелых объектов

Перед тем, как производить выделение объектов на бинаризованных изображениях, необходимо отделить от них нецелые клетки, то есть те, которые попадают на края изображения и не представляют интереса для анализа. Для этих целей был выбран рекурсивный алгоритм, последовательно сканирующий входной массив яркостей и при обнаружении заданного значения осуществляющий заполнение смежных точек с той же яркостью. Алгоритм реализован в универсальном виде и поэтому применим как к массиву люменов, так и к массиву «толстых» стенок, что демонстрируют рисунок 21 и рисунок 22.



а)



б)

Рисунок 21 - Результат удаления нецелых объектов для массива люменов
(а - до удаления, б - после удаления)

Исходный код головной процедуры алгоритма, производящей сканирование массива:

```
procedure TAnalyzePicThread.DeleteCuttedObj (var byte2arr: TByte2Arr;
  const col: Byte);
var
  i, cntr: Integer;
begin
  cntr := ap_PicHeight - 1;

  for i := 0 to cntr do
  begin
    PaintObj(byte2arr, col, 0, i, ap_PicWidth - 1, ap_PicHeight - 1);
    PaintObj(byte2arr, col, ap_PicWidth - 1, i, ap_PicWidth - 1,
      ap_PicHeight - 1);
  end;

  cntr := ap_PicWidth - 1;
  for i := 0 to cntr do
  begin
    PaintObj(byte2arr, col, i, 0, ap_PicWidth - 1, ap_PicHeight - 1);
    PaintObj(byte2arr, col, i, ap_PicHeight - 1, ap_PicWidth - 1,
      ap_PicHeight - 1);
  end;
end;
```

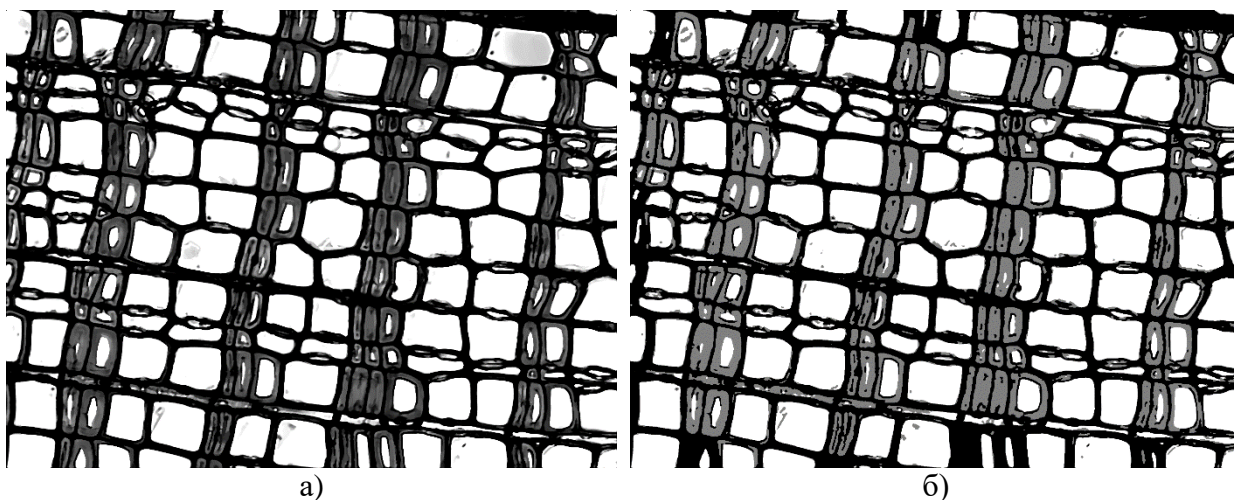


Рисунок 22 - Результат удаления нецелых объектов для массива «толстых» стенок (а - до удаления, б - после удаления)

Исходный код рекурсивной процедуры заполнения найденной области:

```
procedure TAnalyzePicThread.PaintObj(var byte2arr: TByte2Arr; const
col: Byte;
  x, y, w, h: Integer);
begin
  if byte2arr[x, y] = col then
  begin
    byte2arr[x, y] := 0;

    if x = 0 then
      PaintObj(byte2arr, col, x + 1, y, w, h)
    else if x = w then
      PaintObj(byte2arr, col, x - 1, y, w, h)
    else if (x > 0) and (x < w) then
      begin
        PaintObj(byte2arr, col, x - 1, y, w, h);
        PaintObj(byte2arr, col, x + 1, y, w, h);
      end;

    if y = 0 then
      PaintObj(byte2arr, col, x, y + 1, w, h)
    else if y = h then
      PaintObj(byte2arr, col, x, y - 1, w, h)
    else if (y > 0) and (y < h) then
      begin
        PaintObj(byte2arr, col, x, y - 1, w, h);
        PaintObj(byte2arr, col, x, y + 1, w, h);
      end;
    end;
  end;
end;
```

4.3.6 Выделение объектов

Следующим этапом в работе системы является выделение объектов во входных данных, то есть непосредственно сам процесс распознавания. Для


```
TObj = record
    c_Color: Word; // цвет объекта (вспомогательное поле)
    c_MassCenter: TPoint; // центр масс
    c_Arr: array of TPoint; // массив точек объекта
    c_Perimeter: array of TPoint; // массив точек периметра объекта
end;
```

Исходный код головной процедуры обнаружения точек с заданной яркостью и создания элементов массива объектов:

Исходный код рекурсивного алгоритма заполнения массива точек объекта:

```

procedure TAnalyzePicThread.Fill(const byte2arr: TByte2Arr; var obj:
TObj;
    x, y: Integer; const w: Integer; const h: Integer; const col: Byte;
    const fillCol: Byte);
begin
    if byte2arr[x, y] = col then
        begin
            byte2arr[x, y] := fillCol;
        end
    end
end;

```

```

SetLength(obj.c_Arr, Length(obj.c_Arr) + 1);
obj.c_Arr[High(obj.c_Arr)].x := x;
obj.c_Arr[High(obj.c_Arr)].y := y;

if (x > 0) and (x < w) then
begin
  Fill(byte2arr, obj, x - 1, y, w, h, col, fillCol);
  Fill(byte2arr, obj, x + 1, y, w, h, col, fillCol);
end;

if (y > 0) and (y < h) then
begin
  Fill(byte2arr, obj, x, y - 1, w, h, col, fillCol);
  Fill(byte2arr, obj, x, y + 1, w, h, col, fillCol);
end;
end;
end;

```

Результатами работы на данном этапе являются два массива объектов - массив люменов и массив «толстых стенок». При этом длина массива точек каждого объекта - это, по факту, его вычисленная площадь. Математически это можно представить в виде формулы (21):

$$S = \sum_{x=0}^n \sum_{y=0}^m I(x, y), \quad (21)$$

Рисунок 23 демонстрирует отображение массива люменов после работы алгоритма (выделенные объекты закрашены серым, фон - результат этапа удаления нецелых объектов), а рисунок 24 - массива «толстых стенок» (закрашены красным, фон - исходное изображение).

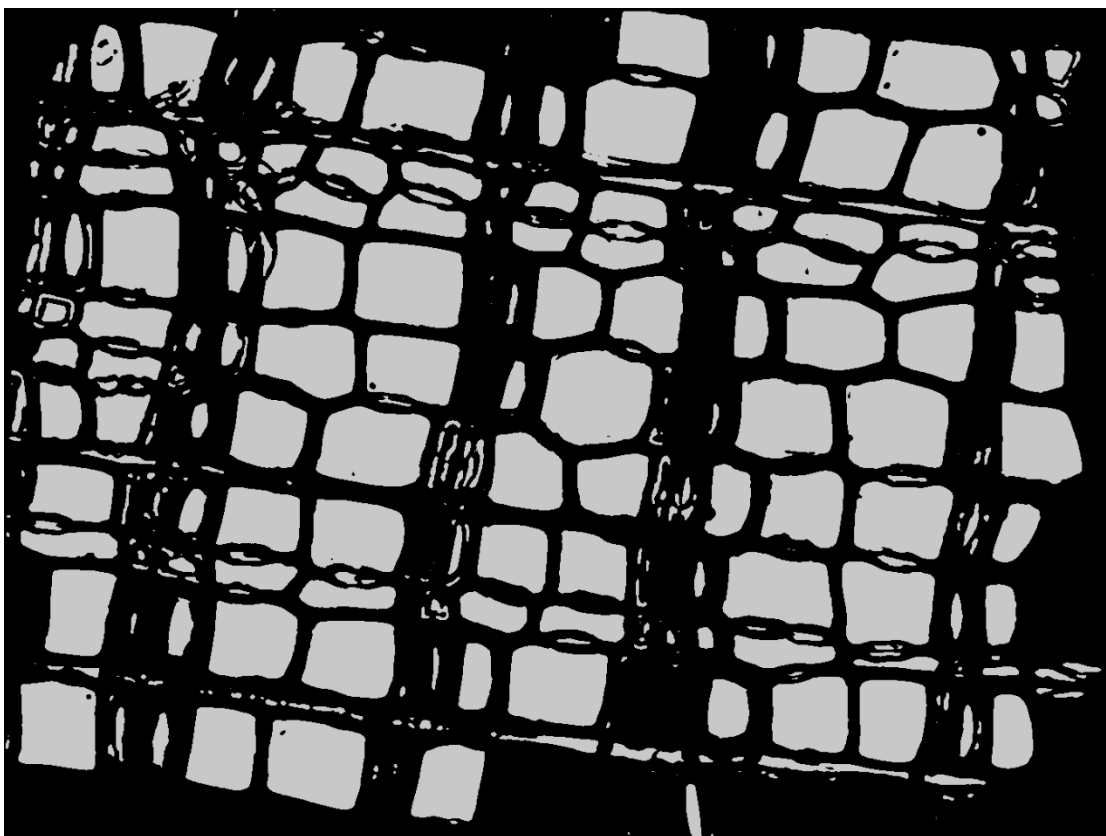


Рисунок 23 - Вид массива выделенных люменов

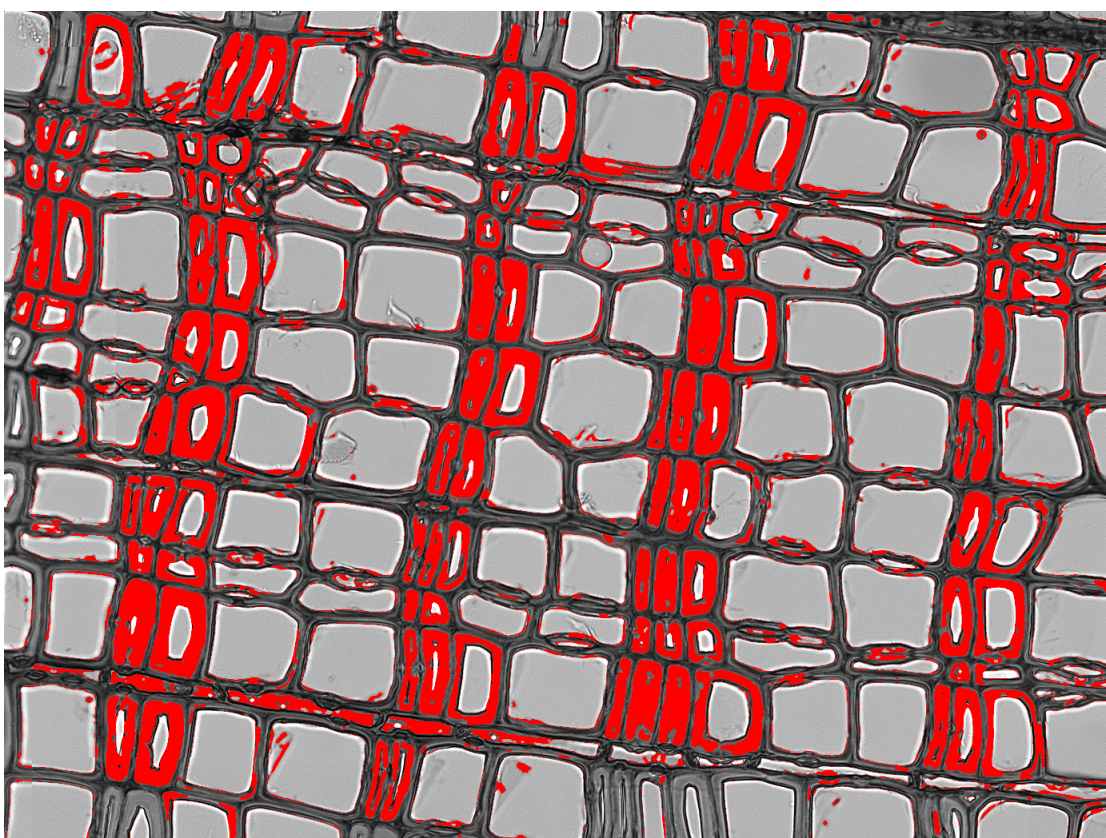


Рисунок 24 - Вид массива выделенных «толстых» стенок

Изм.	Колич.	Лист.	№ док	Подпись	Дата

ДП-230102.65-001019120ПЗ

Лист

59

4.3.7 Морфологическое закрытие объектов

В качестве одного из способов предобработки было применено морфологическое закрытие объектов, которое, однако, в текущей версии не применяется по причине искажения границ объектов. Для реализации закрытия были созданы 3 процедуры: Closing, Erosion и Dilatation.

Процедура Closing представляет собой последовательный вызов процедур дилатации и эрозии объектов. При смене порядка вызова можно аналогично реализовать морфологическое открытие.

Процедуры дилатации и эрозии практически идентичны и в качестве примера достаточно привести исходный код одной из них:

```
procedure TAnalyzePicThread.Dilatation(var inArr: TByte2Arr; const a:
Byte);
var
  y, x, i, j, ii, jj, cntrX, cntrY: Integer;
  tmpArr: TByte2Arr;
begin
  SetLength(tmpArr, ap_PicWidth, ap_PicHeight);

  cntrX := ap_PicWidth - 1;
  cntrY := ap_PicHeight - 1;

  for y := a to cntrY - a do
    for x := a to cntrX - a do
      if inArr[x, y] = 255 then
        // Накладывание аппертуры
        for j := y - a to y + a do
          for i := x - a to x + a do
            // Если хоть один пиксел чёрный
            if inArr[i, j] = 0 then
              begin
                // Заливка области аппертуры белым
                for jj := y - a to y + a do
                  for ii := x - a to x + a do
                    tmpArr[ii, jj] := 255;
                  break;
                end
              end
            else
              tmpArr[x, y] := 255
            end
          else
            tmpArr[x, y] := 0;
          end
        for y := 0 to cntrY do
          for x := 0 to cntrX do
            inArr[x, y] := tmpArr[x, y];
          end
        end;
```

Процедура эрозии полностью аналогична приведённой выше, за исключением того, что заливка области апертуры происходит цветом фона, то есть чёрным.

4.4 Обработка массивов распознанных объектов

4.4.1 Вычисление центров масс объектов

Для облегчения последующей обработки массивов распознанных объектов введён условный параметр «центр масс», являющийся, по сути, средним арифметическим координат всех точек объекта. Математически вычисление центра масс можно выразить формулами (22) и (23):

$$\bar{x} = \frac{\sum_{x=0}^m \sum_{y=0}^n xI(x, y)}{S} \quad (22)$$

$$\bar{y} = \frac{\sum_{x=0}^m \sum_{y=0}^n yI(x, y)}{S} \quad (23)$$

Данный параметр не выходной характеристикой клеточных структур и применяется исключительно для промежуточных вспомогательных операций обработки массивов обнаруженных объектов. Вычисление центра масс происходит в цикле для каждого из массивов по следующему алгоритму:

```
var
  cntr, i, x, y: Integer;
begin
  x := 0;
  y := 0;
  cntr := High(obj.c_Arr);
  for i := 0 to cntr do
    begin
      inc(x, obj.c_Arr[i].x);
      inc(y, obj.c_Arr[i].y);
    end;
  obj.c_MassCenter.x := Round(x / (cntr + 1));
  obj.c_MassCenter.y := Round(y / (cntr + 1));
end;
```

Рисунок 25 иллюстрирует результат вычисления центров масс объектов. Наличие данной характеристики позволяет отсеять ложные объекты, у которых центр масс расположен вне объекта или на его периметре. В данных случаях можно с уверенностью утверждать, что распознанный объект не является клеткой.

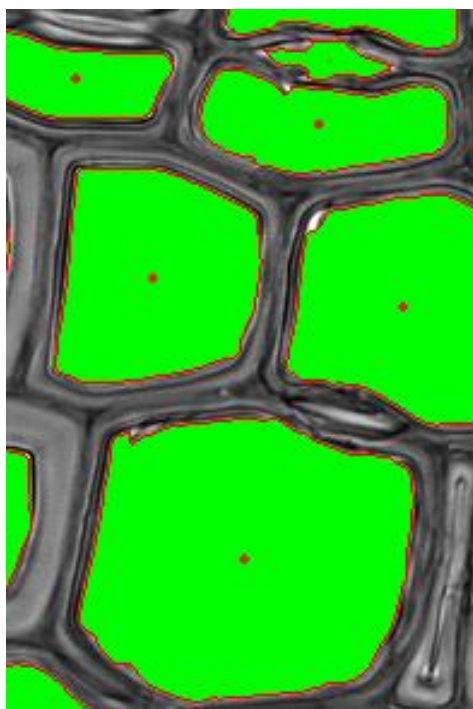


Рисунок 25 - Результат вычисления центра масс объектов

4.4.2 Определение периметра объектов

Ещё одной вспомогательной характеристикой помимо центра масс был выбран периметр объекта, который, в отличие от центра масс, в принципе может выступать в качестве анализируемого параметра. Для нахождения периметра применена операция внутреннего оконтуривания, заключающаяся в следующем: точка принадлежит периметру объекта, если она сама принадлежит объекту, и хотя бы одна из её соседних точек объекту не принадлежит. Алгоритм реализован циклически для каждого из массивов объектов. Результат работы алгоритма демонстрирует рисунок 26. Процедура поиска периметра текущего объекта представлена ниже:

```
procedure TAnalyzePicThread.GetPerimeter(var obj: TObj);
var
  cntr, i { , minX, maxX, minY, maxY } : Integer;
  p, tempP: TPoint;
begin
  SetLength(obj.c_Perimeter, 0);

  cntr := High(obj.c_Arr);
  for i := 0 to cntr do
  begin
    p.x := obj.c_Arr[i].x;
    p.y := obj.c_Arr[i].y;

    tempP.x := p.x;
    tempP.y := p.y - 1;
    // Проверка точки сверху
```

```

if PointInArray(tempP, obj) then
begin
    tempP.x := p.x;
    tempP.y := p.y + 1;
    // Проверка точки снизу
    if PointInArray(tempP, obj) then
    begin
        tempP.x := p.x - 1;
        tempP.y := p.y;
        // Проверка точки слева
        if PointInArray(tempP, obj) then
        begin
            tempP.x := p.x + 1;
            tempP.y := p.y;
            // Проверка точки справа
            if PointInArray(tempP, obj) then
                continue;
        end;
    end;
end;
SetLength(obj.c_Perimeter, Length(obj.c_Perimeter) + 1);
obj.c_Perimeter[High(obj.c_Perimeter)].x := p.x;
obj.c_Perimeter[High(obj.c_Perimeter)].y := p.y;
end;
end;

```

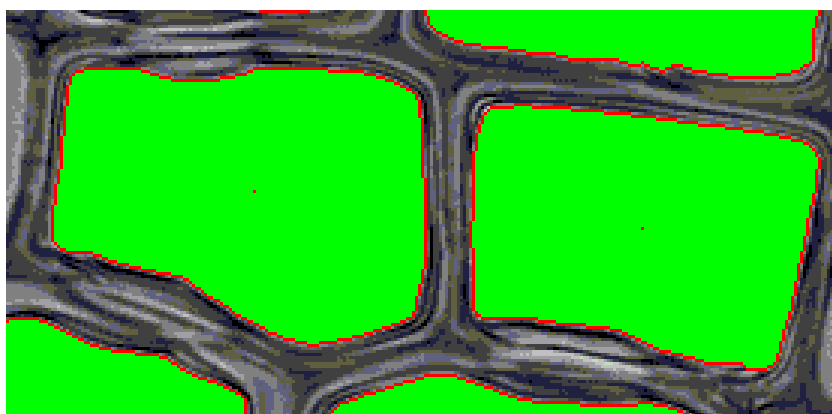


Рисунок 26 - Результат вычисления периметра объекта

4.4.3 Исключение ложных объектов

После обнаружения объектов в массиве входных данных встаёт задача исключить из получившегося множества те элементы, которые не являются клетками, то есть ложные объекты. Решение данной задачи реализовано в несколько этапов:

- удаление мелких объектов;
- удаление люменов с центром масс вне периметра;
- удаление объектов с центром масс на периметре;
- удаление пересечений «толстых» стенок с люменами.

Изм.	Колич.	Лист.	№ док	Подпись	Дата

ДП-230102.65-001019120ПЗ

Лист

63

Задача удаления мелких объектов реализована в процедуре DeleteSmallObj:

```
procedure TAnalyzePicThread.DeleteSmallObj(var objArr: TObjArr; p:
Byte);
var
  i, treshold, curArea: Integer;
begin
  // Вычисление порога
  treshold := Round(GetObjMaxArea(objArr) / 100 * p);

  // Удаление объектов с площадью меньше порога
  i := 0;
  while i < High(objArr) do
  begin
    curArea := Length(objArr[i].c_Arr);
    if curArea < treshold then
      DeleteObjArrElem(objArr, i)
    else
      inc(i);
    end;
  end;
end;
```

Пороговые значения p , ниже которых происходит удаление, выбраны экспериментальным путём индивидуально для массива люменов и массива «толстых» стенок. Для люменов оно равно 2% от максимальной площади объекта массива, для «толстых» стенок - 10%. Одно из возможных направлений дальнейшего развития системы - поиск алгоритма нахождения адаптивного порога удаления. Функция нахождения максимальной площади объекта массива GetObjMaxArea, реализующая циклический обход входного массива тривиальна и представлена ниже:

```
function TAnalyzePicThread.GetObjMaxArea(const objArr: TObjArr):
Integer;
var
  cntr, i, curArea: Integer;
begin
  // Берём за дефолтное значение площадь первого объекта
  result := Length(objArr[0].c_Arr);

  cntr := High(objArr);
  // Сравниваем его с остальными элементами
  for i := 1 to cntr do
  begin
    curArea := Length(objArr[i].c_Arr);
    // Если площадь текущего объекта больше - записываем её в
    результат
    if curArea > result then
      result := curArea;
    end;
  end;
end;
```


Рисунок 27 демонстрирует результат работы алгоритма на примере массива «толстых» стенок. Наглядно видно, что большая часть ложных объектов оказывается исключена из последующей обработки.

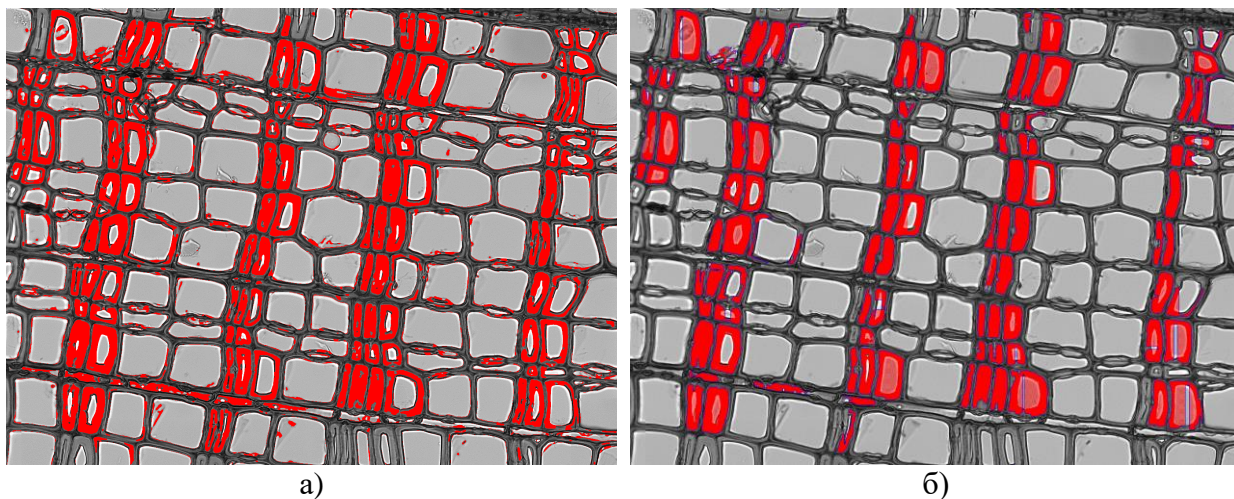


Рисунок 27 - Результат удаления мелких объектов для массива «толстых» стенок (а - до удаления, б - после удаления)

Задачу удаления люменов с центром масс вне периметра объекта реализует процедура DelOutMassCenters:

```
procedure TAnalyzePicThread.DelOutMassCenters(var objArr: TObjArr);
var
  i, minX, maxX, minY, maxY: Integer;
begin
  FillObjMassCenters(objArr);

  i := 0;
  while i < High(objArr) do
  begin
    minX := GetMinX(objArr[i]);
    minY := GetMinY(objArr[i]);
    maxX := GetMaxX(objArr[i]);
    maxY := GetMaxY(objArr[i]);
    if (GetLeftPerimeterPoint(objArr[i], objArr[i].c_MassCenter,
minX)) and
      (GetRightPerimeterPoint(objArr[i], objArr[i].c_MassCenter,
maxX)) and
      (GetTopPerimeterPoint(objArr[i], objArr[i].c_MassCenter, minY))
and
      (GetBottomPerimeterPoint(objArr[i], objArr[i].c_MassCenter,
maxY)) then
      inc(i)
    else
      DeleteObjArrElem(objArr, i);
    end;
  end;
```

Алгоритм процедуры следующий:

- находятся минимальные и максимальные координаты x и y текущего объекта для последующего ограничения зоны поиска;
- последовательно ищутся точки периметра слева, справа, сверху и снизу от центра масс;
- если хоть одна точка не найдена - объект удаляется.

Важной особенностью данного этапа является то, что простая проверка на вхождение центра масс в массив точек объекта приводит к ложным срабатываниям, в силу наличия объектов с пустотами из-за не совсем корректной бинаризации. Если бы центр масс находился внутри такой полости, то объект был бы удалён. Поэтому в ходе разработки было принято решение определять вхождение центра масс в объект именно поиском точек периметра в четырёх направлениях. Такая реализация гарантирует корректные результаты этапа, которые иллюстрирует рисунок 28.

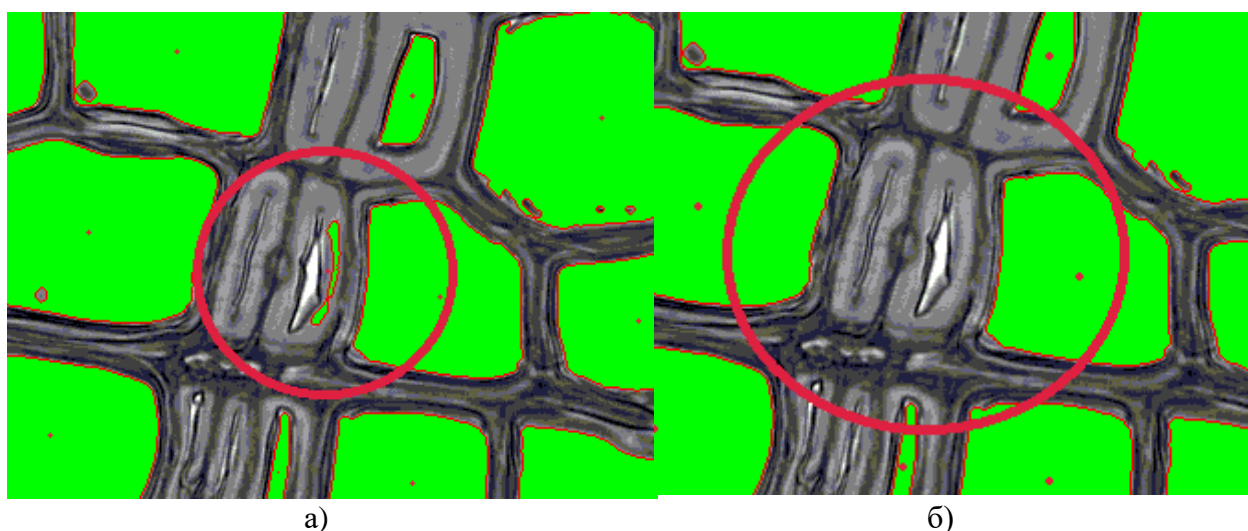


Рисунок 28 - Результат удаления объектов с центром масс вне периметра для массива люменов (а - до удаления, б - после удаления)

Задачу удаления объектов с центром масс на периметре решают процедуры `DeleteMassCenterInPerimeter` и `PointInPerimeterArr`. Первая осуществляет последовательный обход массива объектов, передачу центра масс текущего объекта в процедуру анализа размещения точки на периметре и при положительном результате - удаление объекта. Исходный код процедуры представлен ниже:

```
procedure TAnalyzePicThread.DeleteMassCenterInPerimeter(var objArr:
TObjArr);
var
  i: Integer;
begin
  FillObjMassCenters(objArr);
  FillObjPerimeters(objArr);

  i := 0;
  while i < High(objArr) do
```

```

begin
  if PointInPerimeterArr(objArr[i].c_MassCenter, objArr[i]) then
    DeleteObjArrElem(objArr, i)
  else
    inc(i);
  end;
end;

```

Функция анализа размещения центра масс на периметре проверяет вхождение переданных координат точки в массив точек объекта (в данном случае - массив точек периметра объекта). Если совпадение найдено, функция возвращает значение true, иначе - false:

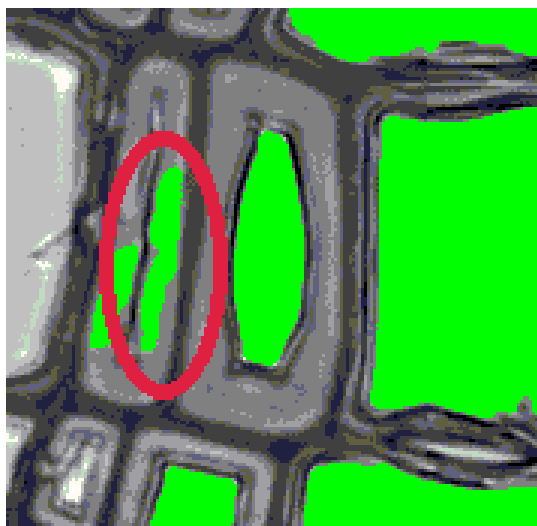
```

function TAnalyzePicThread.PointInPerimeterArr(const p: TPoint;
  const obj: TObj): Boolean;
var
  cntr, i: Integer;
begin
  result := false;

  cntr := High(obj.c_Perimeter);
  for i := 0 to cntr do
    if (p.x = obj.c_Perimeter[i].x) and (p.y = obj.c_Perimeter[i].y)
  then
    begin
      result := true;
      exit;
    end;
  end;
end;

```

Рисунок 29 иллюстрирует результаты работы алгоритма:



а)



б)

Рисунок 29 - Результат удаления объектов с центром масс на периметре для массива люменов (а - до удаления, б - после удаления)

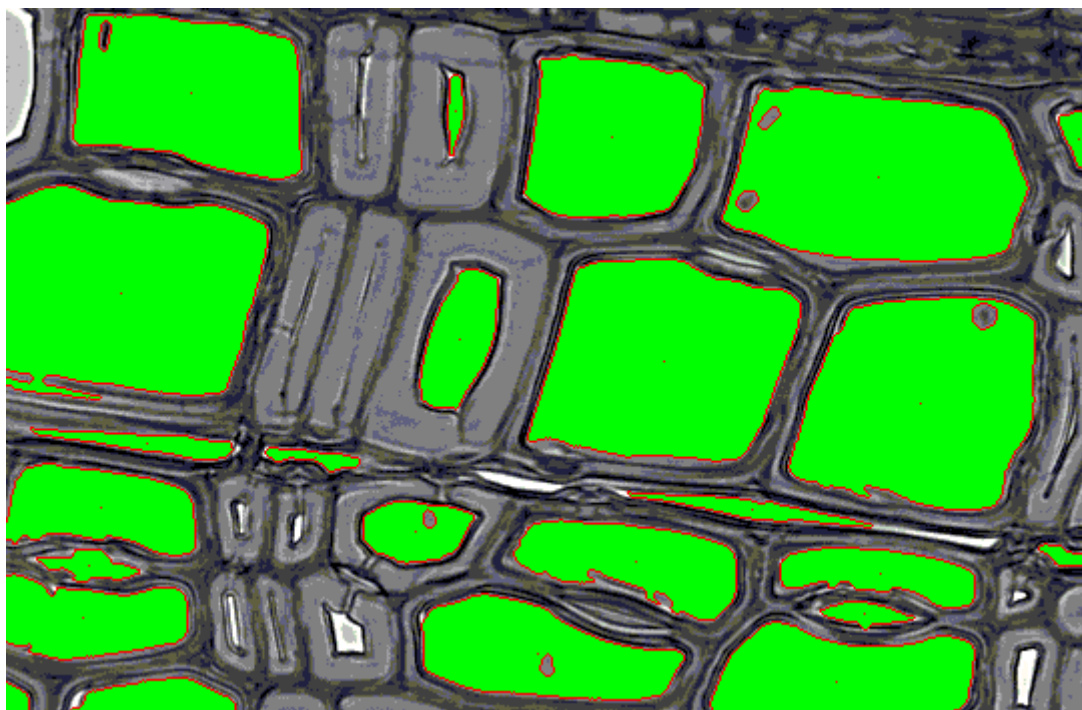
4.4.4 Заполнение полостей объектов

Полученные после выполнения предыдущих этапов объекты содержат в ряде случаев дефекты - внутренние полости, которые образовались по причине погрешностей предобработки. За заполнение пустот внутри объектов отвечает функция FillCurObjEmptyPoints, вызываемая из цикла для каждого текущего объекта. Её алгоритм находит максимальные и минимальные координаты массива точек текущего объекта, а затем последовательно сканирует все точки в получившемся прямоугольнике. Если точка не принадлежит объекту, то проверяется наличие одновременно точек периметра сверху, снизу, справа и слева от неё. Если таковые имеются, то текущая точка добавляется в массив. Результат работы алгоритма демонстрирует рисунок 30. Исходный код функции:

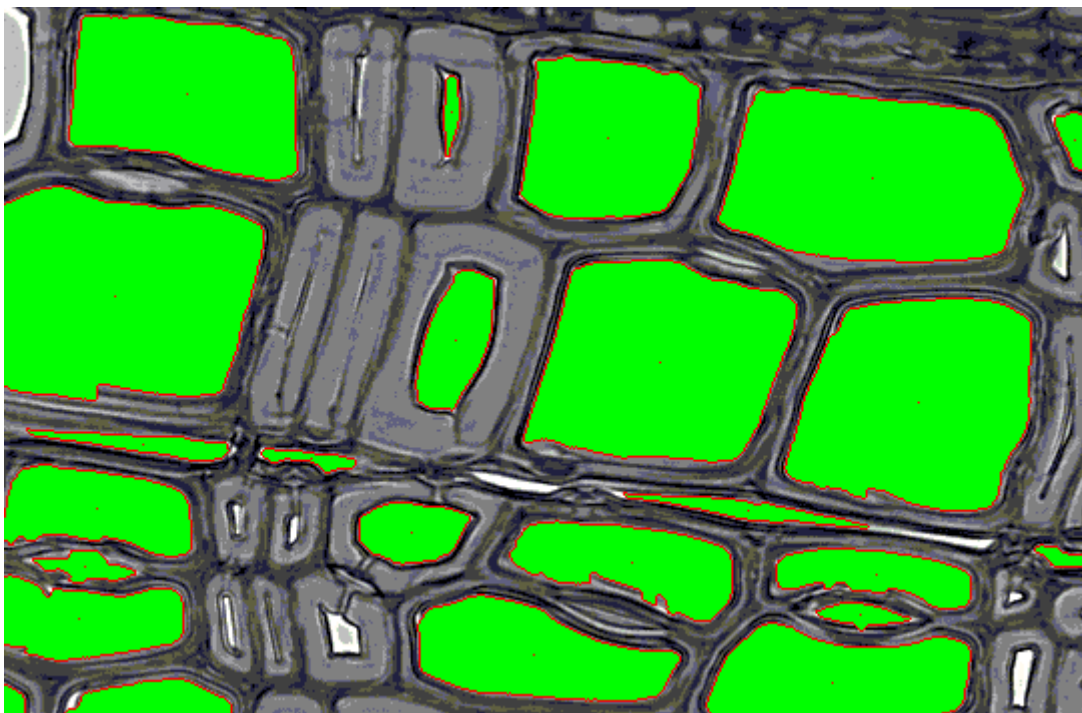
```
function TAnalyzePicThread.FillCurObjEmptyPoints(var obj: TObj):
Boolean;
var
  x, y, minX, maxX, minY, maxY: Integer;
  p: TPoint;
begin
  result := true;

  minX := GetMinX(obj);
  minY := GetMinY(obj);
  maxX := GetMaxX(obj);
  maxY := GetMaxY(obj);

  for x := minX to maxX do
    for y := minY to maxY do
      begin
        p.x := x;
        p.y := y;
        // Если текущая точка не в объекте
        if PointInArray(p, obj) = false then
          begin
            if GetTopPerimeterPoint(obj, p, minY) then
              if GetBottomPerimeterPoint(obj, p, maxY) then
                if GetRightPerimeterPoint(obj, p, maxX) then
                  if GetLeftPerimeterPoint(obj, p, minX) then
                    begin
                      // Добавление текущей точки в объект
                      SetLength(obj.c_Arr, Length(obj.c_Arr) + 1);
                      obj.c_Arr[High(obj.c_Arr)].x := x;
                      obj.c_Arr[High(obj.c_Arr)].y := y;
                      result := false;
                    end;
                  end;
                end;
              end;
            end;
          end;
        end;
      end;
    end;
  end;
```



а)



б)

Рисунок 30 - Результат заполнения пустот внутри объектов для массива люменов (а - до заполнения, б - после заполнения)

4.4.5 Исключение пересекающихся точек объектов

Наконец, когда ложные объекты удалены и внутренние полости заполнены, возникает проблема пересечений «толстых» стенок с люменами в силу погрешностей распознавания. Очевидно, что на этом этапе для решения

Изм.	Колич.	Лист.	№ док	Подпись	Дата

ДП-230102.65-001019120ПЗ

Лист

69

указанной проблемы необходимо вычесть один массив из другого. Для этого разработана процедура DeleteWideWallsOnLumen. На текущий момент работа данного алгоритма - самый длительный этап обработки входного изображения и это всерьёз заставляет задуматься о поиске более быстрого механизма удаления пересечений без прямого перебора «каждый с каждым». Кроме того, после выполнения основной части кода процедуры необходимо повторно вызвать на исполнение алгоритмы удаления мелких объектов и удаления объектов с центром масс на периметре, так как их появление возможно после исключения пересекающихся точек. Результат работы алгоритма иллюстрирует рисунок 31. Исходный код процедуры DeleteWideWallsOnLumen:

```
procedure TAnalyzePicThread.DeleteWideWallsOnLumen(var wideWallArr:
TObjArr; const lumenArr: TObjArr);
var
  i, j, k, cntrI, cntrJ, cntrK: Integer;
begin
  cntrI := High(wideWallArr);
  for i := 0 to cntrI do
    begin
      cntrJ := High(wideWallArr[i].c_Arr);
      for j := 0 to cntrJ do
        if PointInObjArray(wideWallArr[i].c_Arr[j], lumenArr) then
          with wideWallArr[i] do
            begin
              cntrK := High(c_Arr) - 1;
              // Смещение всех последующих элементов массива на 1 позицию
              for k := j to cntrK do
                c_Arr[k] := c_Arr[k + 1];
              // Удаление последнего элемента массива
              SetLength(c_Arr, Length(c_Arr) - 1);
              // // Запись последнего элемента массива на место удаляемого
              // c_Arr[j] := c_Arr[high(c_Arr)];
              // // Удаление последнего элемента массива
              // SetLength(c_Arr, Length(c_Arr) - 1);
            end;
          end;
        end;
      DeleteSmallObj(wideWallArr, 10);
      DeleteMassCenterInPerimeter(wideWallArr);
    end;
  end;
```

По окончании работы процедуры производится повторное удаление мелких объектов и объектов с центром масс на периметре.

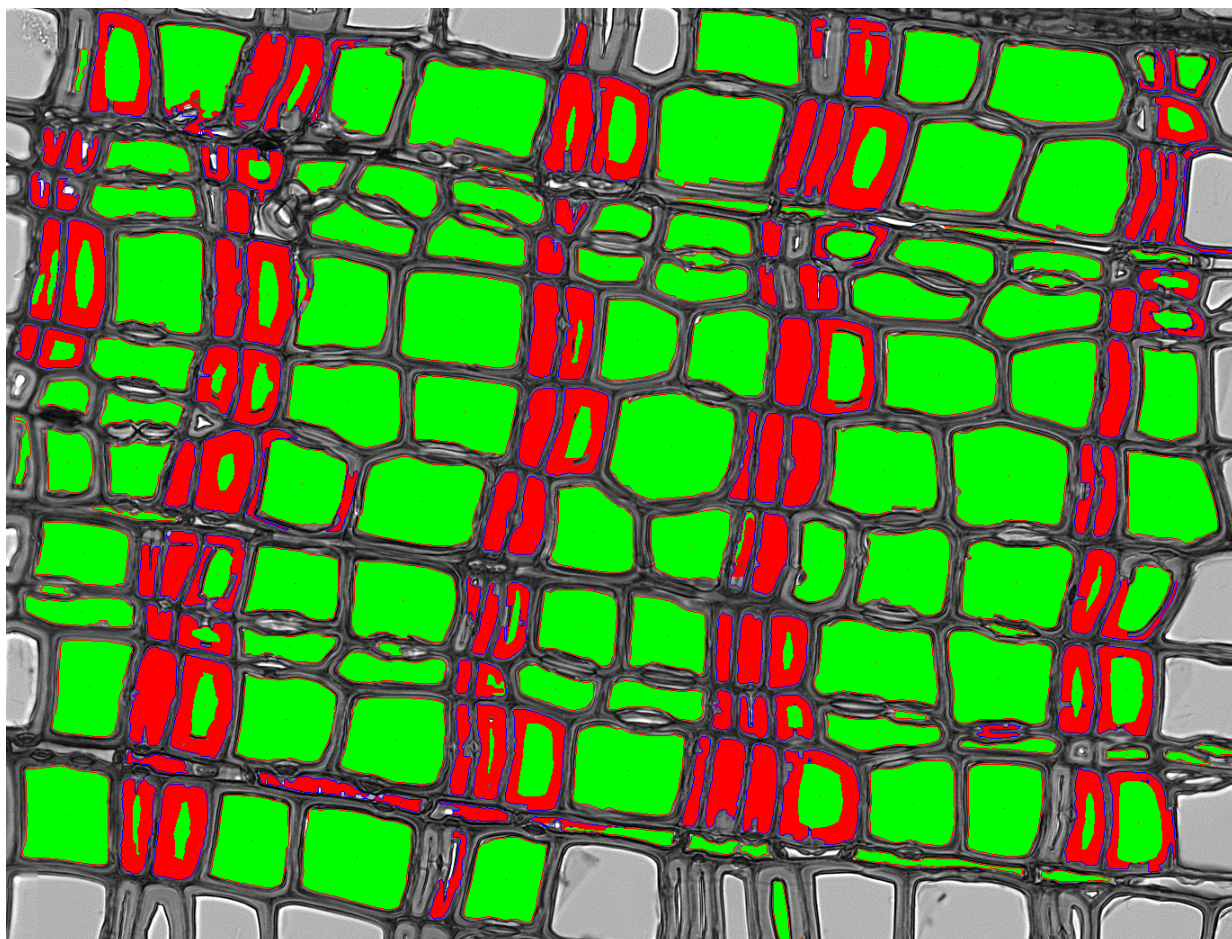


Рисунок 31 - Результат исключения пересечений «толстых» стенок с люменами

4.4.6 Вычисление угла ориентации главной оси изображения

Для последующего вычисления тангенциального и радиального размеров, а также толщин стенок необходимо определить углы наклона объектов. В ходе проектирования системы было решено производить определение среднего угла для всех объектов, на основе которого и производить описанные вычисления. Для этих целей были созданы процедура GetPicOrient, и функции CalcOrientAngle и CalcInertMoment. Результат вычисления угла ориентации главной оси изображения демонстрирует рисунок 32 (белая линия).

Изм.	Колич.	Лист.	№ док	Подпись	Дата

ДП-230102.65-001019120ПЗ

Лист

71

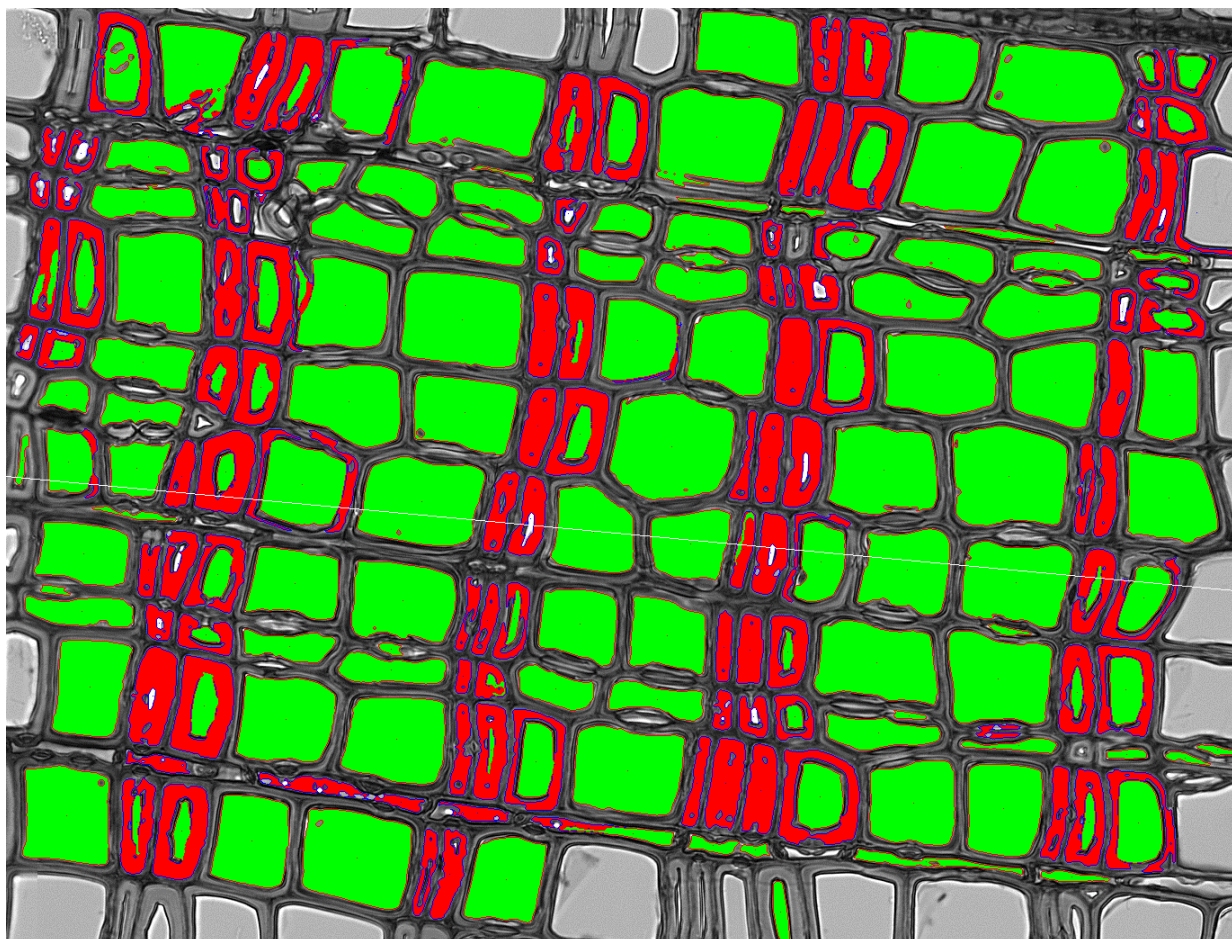


Рисунок 32 - Результат вычисления угла ориентации главной оси изображения (часть операции предобработки исключена для ускорения процесса)

4.5 Выходные результаты работы системы

По завершению всех этапов обработки и анализа изображения полученная информация сохраняется в электронной таблице формата MS Excel. Данный функционал реализован в процедуре SaveAsXLS. Исходный код процедуры представлен ниже:

```
procedure TAnalyzePicThread.SaveAsXLS;
const
  xlWBATWorksheet = -4167;
var
  excelApp, Sheet, data: OLEVariant;
  i, cntr: Integer;
begin
  // Подготовка данных
  cntr := Length(ap_CellArr) + 1;
  data := VarArrayCreate([1, cntr, 1, 10], varVariant);
  // Заполнение заголовков колонок
  data[1, 1] := '№';
  data[1, 2] := 'Тип';
```

Изм.	Колич.	Лист.	№ док	Подпись	Дата

ДП-230102.65-001019120ПЗ

Лист

72


```

...
data[1, 10] := 'Правая стенка';
// Заполнение данных
cntr := High(ap_CellArr);
for i := 0 to cntr do
  with ap_CellArr[i] do
    begin
      data[i + 2, 1] := i + 1;
      data[i + 2, 2] := co_Type; // тип
      ...
      data[i + 2, 8] := co_RightWW; // толщина правой стенки
    end;

// Создание Excel-OLE Object
excelApp := CreateOleObject('Excel.Application');
try
  // Скрытие Excel
  excelApp.Visible := false;
  // Добавление новой рабочей книги
  excelApp.Workbooks.Add(xlWBATWorksheet);
  Sheet := excelApp.Workbooks[1].WorkSheets[1];
  // Sheet.Name := ASheetName;
  // Заполнение листа
  Sheet.Range[RefToCell(1, 1), RefToCell(Length(ap_CellArr) + 1, 8)
  ].Value := data;
  // Сохранение листа
  try
    excelApp.Workbooks[1].SaveAs(ap_FileName + '.xls');
  except
    // Если ошибка
  end;
finally
  // Выход из Excel
  if not VarIsEmpty(excelApp) then
    begin
      excelApp.DisplayAlerts := false;
      excelApp.Quit;
      excelApp := Unassigned;
      Sheet := Unassigned;
    end;
end;
end;

```

Процедура допускает произвольное изменение состава полей путём корректировки содержания переменной data типа OLEVariant.

На текущий момент в проекте реализовано вычисление следующих характеристик:

- площадь люмена;
- периметр люмена;
- радиальный размер люмена;
- тангенциальный размер люмена;
- толщины правой, левой, верхней и нижней стенок.

ЗАКЛЮЧЕНИЕ

Цель выполнения представленного дипломного проекта заключалась в разработке автоматизированной информационной системы параметрического анализа клеточных структур деревьев хвойных пород.

Для достижения указанной цели перед работой был поставлен ряд задач: анализ предметной области, изучение существующих решений в области анализа клеточных структур, изучение методов обработки и распознавания изображений, разработка информационной системы автоматизированного параметрического анализа деревьев хвойных пород.

При решении задачи анализа предметной области были изучены клеточные структуры хвойных, особенности и методика их параметрического анализа.

Для решения задачи изучения существующих решений в области анализа клеточных структур в работе сделан обзор имеющихся наработок в сфере цитологических исследований крови, как одной из самых динамично развивающихся, рассмотрены системы, созданные для этих целей, а также изучен опыт автоматизации параметрического анализа годичных колец хвойных в Сибирском федеральном университете.

Во время решения задачи изучения методов обработки и распознавания изображений был проведён анализ основных методов предварительной обработки и распознавания изображений, потенциально пригодных для использования в настоящем проекте.

При решении задачи разработки информационной системы автоматизированного параметрического анализа деревьев хвойных пород был разработан и описан прототип АИС на основе проделанной на предшествующих этапах исследовательской работы.

Таким образом, задачи дипломного проекта решены в полном объёме, цель достигнута - разработан прототип автоматизированной информационной системы параметрического анализа клеточных структур деревьев хвойных пород.

Изм.	Колич.	Лист.	№ док	Подпись	Дата

ДП-230102.65-001019120ПЗ

Лист

74

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Быков, Р.Е. Цифровое преобразование изображений: учебное пособие для ВУЗов / Р.Е. Быков. - Москва: Горячая линия, 2012. - 228 с.
2. Гонсалес, Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. - Москва: Техносфера, 2012. - 1104 с.
3. Дружинин, Е.А. Обработка и распознавание изображений в системах автоматизированного проектирования / Е.А. Дружинин, О.К. Погудина, И.Н. Бабак, А.В. Губарев. - Харьков: ХАИ, 2011. - 52 с.
4. Красильников, Н.Н. Цифровая обработка 2D- и 3D-изображений / Н.Н. Красильников. - Санкт-Петербург: БХВ-Петербург, 2012. - 608 с.
5. Мерков, А.Б. Распознавание образов. Введение в методы статистического обучения / А.Б. Мерков. - Москва: Едиториал УРСС, 2011. - 256 с.
6. Силкин, П.П. Методы многопараметрического анализа структуры годичных колец хвойных: монография / П.П. Силкин. - Красноярск: Сибирский федеральный университет, 2010. - 335 с.
7. Фомин, Я.А. Распознавание образов: теория и применения / Я.А. Фомин. - Москва: ФАЗИС, 2012. - 429 с.
8. Статья «Медицинская тематика» [Электронный ресурс]: Открытое акционерное общество «Радиотехнический институт имени академика А.Л. Минца». - Режим доступа: <http://www.rti-mints.ru/medicinskaya-tematika/>
9. Статья «Методы исследования крови», [Электронный ресурс]: Центр Владимира Цыганкова в Санкт-Петербурге. - Режим доступа: <http://alcoholismhls.ru/2015/07/17/metody-issledovaniya-krovi/>

Изм.	Колич.	Лист.	№ док	Подпись	Дата

ДП-230102.65-001019120ПЗ

Лист

75

ПРИЛОЖЕНИЕ А

Техническое задание

Настоящее техническое задание содержит требования к автоматизированной информационной системе параметрического анализа клеточной структуры деревьев хвойных пород, а именно:

- основание для разработки системы;
- назначение разработки системы;
- требования к программе и программному изделию;
- требования к программной документации;
- требования к порядку внесения изменений в техническое задание.

А.1 Основание для разработки

А.1.1 Основание для разработки

Основанием для разработки является задание на выпускную квалификационную работу в форме дипломного проекта.

А.1.2 Исполнитель

Исполнителем разработки, выполняемой на основании данного технического задания, является студент группы ЗКИ 10-08-1 ИКИТ СФУ Клёнкин Алексей Юрьевич.

А.1.3 Наименование

Проекту, разрабатываемому на основании данного технического задания, присваивается наименование: «Информационная система параметрического анализа клеточной структуры деревьев хвойных пород», в дальнейшем по тексту именуемая «информационная система».

А.2 Назначение разработки

Информационная система предназначена для выполнения параметрического анализа клеточной структуры деревьев хвойных пород путём обработки фотоизображений образцов с электронного микроскопа. клеточных структур деревьев хвойных пород.

А.3 Требования к программе и программному изделию

А.3.1 Требования к составу

Информационная система должна состоять из одного модуля, выполняющего все требуемые функции.

А.3.2 Требования к функциональным характеристикам

АИС учёта договоров должна состоять из одного модуля, позволяющего выполнять требуемые функции.

Загрузка изображений. Программа должна предоставлять механизм для загрузки фотоизображений с НМЖД или внешнего носителя и их отображения в графическом интерфейсе.

Нормализация изображений. На стадии подготовки фотоизображений программа должна преобразовывать их в 8-битный формат и переводить в режим 256 оттенков серого. Программа должна автоматически определять необходимость подобных преобразований.

Снижение шума. На стадии подготовки фотоизображений программа должна с помощью предусмотренных в её структуре алгоритмов произвести фильтрацию с целью снижения уровня зашумлённости.

Бинаризация изображений. На стадии подготовки фотоизображений программа должна производить бинаризацию для перевода многоградационных по яркости изображений к двухградационным.

Маркировка. На стадии распознавания клеточной структуры на фотоизображениях программа должна производить маркировку связных областей в результате которой получается изображение, в котором каждому объекту присвоен отдельный номер.

Удаление нецелых объектов. На стадии распознавания клеточной структуры на фотоизображениях программа должна производить удаление нецелых объектов, то есть не полностью вошедших в кадр поля зрения.

Выделение клеточной структуры. На стадии распознавания клеточной структуры на фотоизображениях программа должна производить распознавание и выделение в отдельные смысловые единицы клетки.

Параметрический анализ. На стадии измерений программа должна производить определение следующих геометрических характеристик распознанных объектов:

- радиальный размер люмена - Lu;
- радиальный размер клетки - D;
- тангентальный размер клетки - T;
- толщина двойной клеточной стенки - WW.

Сохранение результатов анализа. Программа должна производить сохранение результатов проведённого параметрического анализа клеточных структур деревьев хвойных пород в отдельный для каждого исходного изображения файл установленного формата на НМЖД или внешний носитель.

А.3.3 Требования надёжности

Надёжность информационной системы должна быть обеспечена правильностью алгоритмических решений и программирования. Информационная система в составе ПО ПЭВМ должна выполнять функциональные задачи по назначению в течении всего срока эксплуатации ПЭВМ. Время восстановления информационной системы после отказа не должно превышать 10 минут.

А.3.4 Условия эксплуатации

Информационная система должна функционировать в соответствии с заданным в настоящем техническом задании требованием в составе ПО ПЭВМ при эксплуатации ПЭВМ. Условия эксплуатации информационной системы должны соответствовать условиям эксплуатации ПЭВМ, требования к которым предъявляются в эксплуатационной документации ПЭВМ или её составных частей.

А.3.5 Требования к составу и параметрам технических средств

Информационная система должна функционировать в составе ПЭВМ со следующими характеристиками:

- процессор Intel Core i5 с тактовой частотой 2.6ГГц и выше;

- объем ОЗУ 4Гб и выше;
- графический адаптер AMD Radeon HD7600MSeries;
- сетевая карта совместимая с Ethernet;
- манипулятор типа «мышь»

Штатным носителем информационной системы является НЖМД ПЭВМ. Технологическим носителем информационной системы является компакт-диск.

А.3.6 Требования к информационной и программной совместимости

Язык и среда программирования информационной системы определяются Исполнителем на этапе разработки.

Информационная система должна функционировать на ПЭВМ с операционной системой не ниже Microsoft Windows 7.

Для реализации интерфейса информационной системы должны использоваться средства графического интерфейса операционной системы.

Для реализации межпрограммного интерфейса должны использоваться средства буферизации операционной системы.

А.3.7 Требования к маркировке и упаковке

Маркировка компакт-диска с информационной системы должна проводиться в соответствии с требованиями ГОСТ 19.102-77 ЕСПД («ЕСПД. Стадии разработки программ и программной документации»).

А.3.8 Требования к транспортировке и хранению

Информационная система должна транспортироваться в составе ПЭВМ, записанной на НЖМД ПЭВМ или на компакт-диск. Условия транспортирования информационной системы в составе ПЭВМ должны соответствовать условиям транспортирования ПЭВМ, требования к которым предъявляются в эксплуатационной документации ПЭВМ или её составных частей.

Условия транспортирования информационной системы на компакт-диске должны соответствовать условиям транспортирования компакт-диска, требования к которым предъявляются в эксплуатационной документации на компакт-диск.

Информационная система должна храниться в составе ПЭВМ, записанной на НЖМД ПЭВМ или на компакт-диск. Условия хранения информационной системы в составе ПЭВМ должны соответствовать условиям хранения ПЭВМ, требования к которым предъявляются в эксплуатационной документации ПЭВМ или её составных частей.

Условия хранения информационной системы на компакт-диске должны соответствовать условиям хранения компакт-диска, требования к которым предъявляются в эксплуатационной документации на компакт-диск.

А.3.9 Специальные требования

Специальные требования не предъявляются.

А.4 Требования к программной документации

А.4.1 Требования к составу документации

Состав документации определяется Исполнителем на этапе разработки перечнем разрабатываемой документов и согласовывается с Заказчиком. В комплект документации в обязательном порядке должны входить:

- пояснительная записка;
- загрузочные модули;

А.4.2 Требования к оформлению документации

Программная документация должна быть оформлена в соответствии с ЕСПД.

Документация, прилагаемая к программному продукту, должна быть разработана и оформлена в соответствии с ЕСКД.

А.5 Порядок внесения изменений в техническое задание

Настоящее техническое задание может уточняться и дополняться в процессе разработки информационной системы по согласованию с Заказчиком.